Department of AERONAUTICS and ASTRONAUTICS
STANFORD UNIVERSITY

D. K. FREDERICK

PIECEWISE-LINEAR SWITCHING FUNCTIONS FOR QUASI-MINIMUM-TIME CONTACTOR CONTROL SYSTEMS

UNPUBLISHED PRELIMINARY DATA

DECEMBER 1963

PREPARED FOR THE NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
UNDER GRANT NsG 133-61

SUDAER NO. 178

Department of Aeronautics and Astronautics
Stanford University
Stanford, California

PIECEWISE-LINEAR SWITCHING FUNCTIONS FOR QUASI-MINIMUM-TIME CONTACTOR CONTROL SYSTEMS

by

D. K. Frederick

SUDAER No. 178
December 1963

This research was sponsored by
the National Aeronautics and Space Administration
under Research Grant NSG-133-61

16738

ABSTRACT

A

Piecewise-linear switching functions are investigated for minimum-time regulator systems represented by a linear, constant parameter plant and a single controller. It is assumed that the plant transfer function has only poles and that the controller is an ideal contactor. Consideration is limited to switching functions which are the sum of linear or piecewise-linear functions of single variables.

A performance criterion is defined in terms of the response times for a number of initial conditions. Optimal linear and piecewise-linear switching functions are found by searching the surface relating the performance criterion and the undetermined switching function parameters. It is shown that the initial conditions used to define the criterion function have a substantial effect upon the smoothness of the performance surface. Methods are given for selecting the initial conditions used to define the criterion function so that the performance surface is amenable to standard surface searching procedures.

A qualitative method based upon root-locus techniques and the existence of certain periodic solutions is given for designing sub-optimal linear switching functions, in order to initiate the search of the performance surface in a region where the surface exists and for determining which components of the switching function should be made piecewise-linear functions.

The design methods are illustrated by synthesizing piecewise-linear switching functions which yield close-to-optimal step-function responses for third- and fourth-order plants having all of their poles on the imaginary axis of the s-plane. These plants were chosen because they are inherently more difficult to control than plants having well-damped poles. The qualitative design procedures are applied also to a third-order plant having well-damped poles.

It is concluded that easily-implemented, piecewise-linear switching functions can provide close-to-optimal response for a wide variety of plants and initial conditions. The design procedures are applicable to plants of high-order, although, for plants of dominant fifth- or higher-order, the design of optimal piecewise-linear switching functions leads to a surface searching problem of considerable complexity. AUTHOR

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF SYMBOLS

$\underline{e}$	the error state-vector
e	the error variable
e_i	a component of $\underline{e}$
k	the index for the initial conditions defining the cost
m	the index for the gradient surface searching procedure
n	the order of the plant
s	the Laplace transform argument
t	the time variable
u	the scalar control variable
w_k	the weighting factor for the k^{th} initial condition
$\underline{x}$	the canonical state-vector
x_i	a component of $\underline{x}$
A	a dimension of the cost-free region S
C_i	a constant value of $\partial\phi/\partial\rho_i$
$\underline{D}$	the distribution matrix of the control variable
E(s)	Laplace transform of e(t)
F	the matrix defining the plant differential equation
I	the minimum value of J for a particular class of switching functions
J	the cost
K	the number of initial conditions used to define J
$K(\sigma)$	an equivalent contactor gain
N	the number of unsuccessful random perturbations taken before V is halved
$\underline{R}$	an arbitrary parameter vector

$\underline{\Delta R}$	a change in $\underline{R}$
S	the cost-free region surrounding the state-space origin
T_k	a response time for the k^{th} initial condition
T_o	the optimal time for the state to reach the origin
T_s	the optimal time for the state to reach S
V	the scale factor of parameter perturbations in the random search procedure
α	the real part of a real switching function zero
α_m	the negative of the maximum real part of all switching function zeros
γ	the gain of the gradient search procedure
δ	the parameter change used to measure gradient of J
ϵ	an arbitrary constant
ζ	the damping ratio of a pair of complex switching-function zeros
θ_i	root-locus angles
ν	$\sqrt{1 - \zeta^2}$
$\underline{\xi}$	the linear switching-function parameter vector
ξ_i	a component of $\underline{\xi}$
$\{\rho\}$	the array of PWL switching-function parameters
ρ_{ij}	the component of $\{\rho\}$ in the i^{th} row and j^{th} column
$\underline{\rho}$	the vector of linear switching-function coefficients
ρ_i	a component of $\underline{\rho}$
σ	the switching function, a scalar
σ_i	a linear or PWL function contained in σ
Φ	the angle of departure of the root locus from a pole
ψ	the quarter-period of the dominant periodic solution

LIST OF SYMBOLS (Cont.)

ω	the natural frequency of a pair of complex switching function zeros
$\Sigma(s)$	the Laplace transform of $\sigma(t)$
T	(superscript) the matrix transpose operator
o	(superscript) the value at $t = 0$
$\cdot$	(superscript) differentiation with respect to t
$\wedge$	(superscript) the optimal value
m	(subscript) the maximum value

ACKNOWLEDGMENT

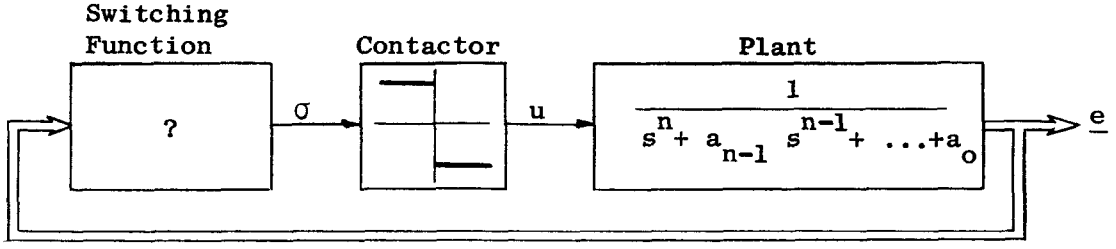
The author would like to express his gratitude to Professor G. F. Franklin for suggesting the topic of this dissertation and for supervising the research and to Professors I. Flügge-Lotz and R. H. Cannon, Jr. for their helpful suggestions in preparing the manuscript for publication. In addition, thanks are due to the author's fellow graduate students for numerous helpful discussions and for assistance in obtaining the hybrid computer results.

This research was supported by the National Aeronautics and Space Administration under Research Grant NSG 133-61, for which support the author is very grateful. The work was performed at the Stanford Electronics Laboratories.

I. INTRODUCTION

A. PROBLEM STATEMENT

The problem under consideration is the design of a switching function for the feedback control system shown below.* The plant, whose



transfer function contains only poles, is described by the n^{th}-order linear constant-coefficient differential equation

$$\dot{\underline{e}} = F\underline{e} + \underline{D}u \qquad (1.1)$$

where $\underline{e}$ is an n x 1 column state vector,

F is a known n x n system matrix,

$\underline{D}$ is a known n x 1 distribution matrix of the control u,

u is a bounded scalar with $|u| \leq 1$,

and the feedback is obtained by having $u = -\text{sgn}\ \sigma(\underline{e})$. The n-component state vector $\underline{e}$ may be the error variable e and its first (n-1) derivatives or the state variables of two or more coupled plants of lower-order. It will be assumed that all components of $\underline{e}$ can be

---

* Double lines denote a vector variable, single lines a scalar.

observed and controlled. The above system can represent either a regulator process with zero input and non-zero initial conditions which are to be reduced to the state-space origin or a control process with an input function having one or more eigenvalues identical to those in the plant, which implies that the input can be imbedded and treated as an initial condition on the error state vector.

The system is essentially deterministic in that it is not subject to disturbances. However, the initial conditions or, equivalently, the imbeddable inputs may be random. The fact that the output of the actuator or controller is bounded implies the existence of a minimum time in which the initial conditions may be reduced to zero and it is this time which will be taken as the criterion of performance. Consideration will be limited to contactor, i.e., bang-bang, actuators for the following reasons:

1. Many actuators inherently operate in this manner.

2. Saturating linear actuators which operate in the saturated mode a large percentage of the time are well approximated by contactors.

3. The optimal control leading to the minimum-time response of the system described by Eq. (1.1) is bang-bang.

The differential equation representing the physical plant has been normalized with respect to the magnitude of the control force so that $|u| = 1$. Assuming that Eq. (1.1) is known, the problem is to determine the switching function $\sigma(\underline{e})$ as a function of the state variables so as to minimize the performance criterion. Equivalently, the problem is to find an $(n-1)$-dimensional switching surface described by $\sigma = 0$ which separates the n-dimensional state-space into two regions, one where $u = +1$, the other where $u = -1$.

B. PREVIOUS WORK

Pontryagin and his co-workers [Ref. 1] have given necessary conditions which an optimal control function $u(t)$ must satisfy under the assumptions described above. However, this optimal control is given in terms of the adjoint variables whose functional form is known but whose

values are unknown. Using the form of the adjoint variables the optimal switching surface can be constructed for some relatively simple plants (second-and third-order) but even then the switching function might have such a complicated dependence upon the state variables as to be useless as an engineering solution.

Therefore, attention has been given to quasi-optimal switching functions which give less than optimal performance but may be more feasible to implement than the true optimal function. Hubbs [Ref. 2] has used a switching surface described by the first few terms of a Taylor-series expansion of the optimal surface about the state-space origin. Although this technique is valid for systems less restrictive than those under consideration here, it is feasible for only second-and third-order plants and works well only when the norm of $\underline{e}$ (denoted by $\|\underline{e}\|$) is small, say $\|\underline{e}\| < |u|$. Schmidt [Refs. 3,4] has treated saturating linear systems subject to step inputs by making the feedback gains of one or more of the state variables be nonlinear functions such that the actuator comes out of saturation at the same time the equivalent optimal contactor system would switch for the first time. However, when the plant has lightly damped oscillatory roots the design method is not applicable. Numerous authors have suggested using the optimal switching surface near the origin and replacing it by linear segments in the more distant regions of the state-space, such as described by Feld'baum [Ref. 5] .

Nonlinear functions useful for plants having lightly damped oscillatory roots and subject to large disturbances, say $\|\underline{e}\| > |u|$, have been considered by Flügge-Lotz and Titus [Ref. 6] , but some other provision must be made for small disturbances. Flügge-Lotz and Lindberg [Ref. 7] used the switching function $\sigma = e + k\,\dot{e} + f(\dot{e})\ddot{e}$ to obtain good step-function response for the plant with transfer function $1/s(s^2 + 1)$ by adjusting $f(\dot{e})$ so that the first switching time was optimal. However, the performance was found to deteriorate rapidly when the initial values of $\dot{e}$ and $\ddot{e}$ were not precisely zero. The simplest possible switching function, first studied by Flügge-Lotz [Ref. 8] , is linear switching, where σ is a linear combination of the states, i.e., $\sigma = \underline{\rho}^T \underline{e}$, where $\underline{\rho}^T$ is the transpose of $\underline{\rho}$. This

relationship corresponds to the switching surface being a single plane in the n-dimensional state-space. If the optimal trajectory from a particular initial condition requires only (n-1) switchings, it is usually* possible for a linear switching function to provide an optimal response for that initial condition and any others requiring the same switching points by passing a plane (or hyperplane) through the (n-1) switching points and the origin. Flügge-Lotz et al. [Refs. 7, 9, 10] have shown that for some parameter combinations of the general third-order plant $1/(s + \alpha)(s^2 + 2\zeta s + 1)$ the values of the optimal linear switching coefficients remain relatively constant for initial conditions along a large portion of the e axis and the linear switching function can yield good performance. However, for many plant parameter values and for more general initial conditions, linear switching is unsatisfactory.

Therefore, one is led to consider piecewise-linear (hereafter abbreviated PWL) switching functions in an effort to provide close-to-optimal response for initial conditions in all or a relatively large portion of the state-space while retaining much of the simplicity of linear switching.

C. OUTLINE OF CHAPTERS

In Chapter II, PWL switching functions are described using a second-order plant as an example. The mathematical formulation given implies the existence of a performance surface which can be searched to find

---

*Necessary and sufficient conditions for linear switching to yield an optimal trajectory requiring (n-1) switchings are that:

1. The switching plane contains the (n-1) switching points and the origin.

2. Those portions of the optimal trajectory corresponding to u = +1 lie on one side of the switching plane and those corresponding to u = -1 lie on the other side.

An example where (2) cannot be satisfied is the plant with transfer function $1/s^2(s^2 + 1)$ and the initial state $(\underline{e}^o)^T = (20, 0, 0, 0)$.

those switching function parameter values which will yield the minimum cost. In Chapter III it is shown how to analyze qualitatively the system performance with a linear switching function in order to obtain a starting point for searching the performance surface. Also the qualitative analysis is extended to PWL switching functions. Chapters IV and V illustrate the design procedure and the type of results which might be expected by synthesizing PWL switching functions for third- and fourth- order plants. The plants considered in detail are described by transfer functions which have all of their poles on the imaginary axis of the s-plane, thereby insuring that they cannot be treated as plants having dominant poles of a lower order. Chapter VI consists of a summary of the results obtained and suggestions for future investigation.

It is shown that PWL switching functions which are easily implementable can give close-to-optimal responses for a variety of plants and initial conditions and both qualitative and quantitative procedures are given for the design of linear and PWL switching functions. It is felt that this work will help to bridge the gap between the fields of optimal control theory and control system design, particularly in the area of satellite attitude and trajectory control.

II. PERFORMANCE SURFACES FOR PWL SWITCHING FUNCTIONS AND THE MINIMUM-TIME CRITERION

A. PWL SWITCHING

Since a general PWL function of n variables can be extremely complicated to generate and analyze and since the main purpose in considering PWL switching functions is to achieve a function which is easily implemented, consideration will be limited to only a very restrictive subclass of PWL functions. This class will consist of those functions which are the summation of up to n symmetric PWL functions of single state variables, where n is the order of the plant. In other words, the switching function σ will be limited to the class

$$\sigma(\underline{e}) = \sum_{i=1}^{n} \sigma_i(e_i) \tag{2.1}$$

where one or more of the $\sigma_i(e_i)$ may be symmetric PWL functions of the type shown in Fig. 1.

Examination of Fig. 1 indicates that ρ_{i1} is the slope of the central portion of σ_i, that ρ_{i2}, ρ_{i4}, etc. are the breakpoints, and that ρ_{i3}, ρ_{i5}, etc. are the slopes after the corresponding breakpoints. If the i^{th} component of σ is not PWL then only ρ_{i1} is defined and ρ_{i2}, ρ_{i3}, ρ_{i4}, ρ_{i5}, etc. are undefined for that particular switching function. Since the output of the switching function σ goes directly to the contactor, σ can be multiplied by any positive constant without affecting $u(t)$ or, equivalently, ρ_{n1}, the linear switching coefficient of e_n, can be set equal to unity. It is convenient to arrange all of the switching function parameters ρ_{ij} in the array $\{\rho\}$ having r elements which are both defined and arbitrary contained in n rows and a number of columns dependent upon the maximum number of breakpoints in any single component σ_i. For example, if the switching function for a third-order plant had its σ_1 component defined as in Fig. 1 and both σ_2 and σ_3 were linear functions, the array of switching function

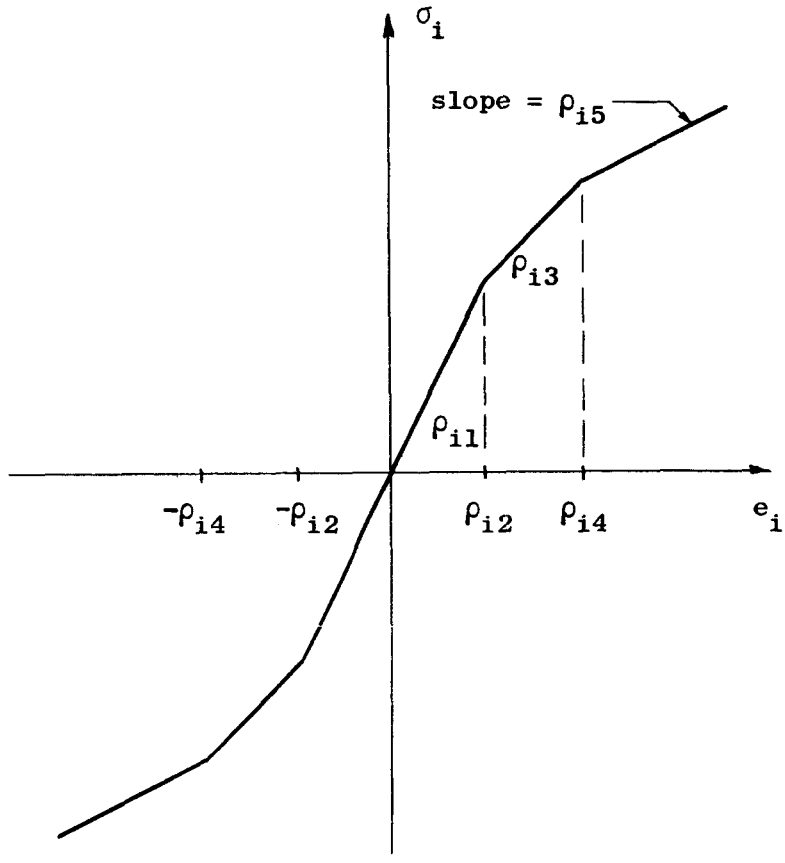


FIG. 1. TYPICAL PWL SWITCHING-FUNCTION COMPONENT

parameters would have $r = 6$ and would appear as

$$\{\rho\} = \begin{Bmatrix} \rho_{11} & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} \\ \rho_{21} & x & x & x & x \\ 1 & x & x & x & x \end{Bmatrix}$$

where the x's denote undefined elements.

It should be emphasized that the array $\{\rho\}$, being merely a collection of numbers, is <u>not</u> a matrix and thus has <u>no</u> algebraic properties.

In the case of linear switching, the switching function may be written as

$$\sigma = \underline{\rho}^T \underline{e}$$

where $\underline{\rho}$ is the parameter vector, i.e., column matrix, whose n components are the linear switching coefficients and it obeys the laws of matrix algebra. If the PWL switching function corresponding to the array $\{\rho\}$ is considered to be a linear switching function represented by the vector $\underline{\rho}$ which was made PWL by the addition of breakpoints and changes of slope, then the vector $\underline{\rho}$ will be the first column of the array $\{\rho\}$, i.e.,

$$\underline{\rho}^T = (\rho_{11}, \rho_{21}, \rho_{31}) .$$

For second-order plants the restriction of $\sigma(\underline{e})$ to the class of functions satisfying Eq. (2.1) involves no loss in generality, as the most general PWL function in two dimensions is a function of a single variable. While it is impossible to state precisely what loss in performance will be incurred for higher-order systems by this restriction, the simplification gained in realizing and analyzing the restricted class of PWL functions makes it a logical area for investigation before proceeding to more complicated types.

So far, no mention has been made of the number of PWL components σ is to have or the number of breakpoints the PWL σ_i are to contain. The emphasis in the work to follow will be upon finding switching functions with a minimal number of PWL elements and the PWL switching function will be considered as a modification of a linear switching function. While no precise statements can be made regarding the minimal number of PWL elements to be used, it will be shown by the examples presented in later chapters that, at least in these instances, a significant improvement in performance can be made over that attainable with linear

switching by introducing only one or two breakpoints. Aside from the resultant simplicity, a PWL switching function with a minimal number of elements is of obvious use in an adaptive system where the switching function is to be adjusted during operation.

B. SPECIFICATION OF THE PERFORMANCE SURFACE

In order to evaluate quantitatively a particular switching function it is necessary to define a scalar performance criterion, referred to as the cost, whose value will depend upon the switching function parameters. The cost (denoted by J) and the r arbitrary elements of $\{\rho\}$ define a performance surface in an $(r+1)$-dimensional space and the desired switching function corresponds to that value of $\{\rho\}$ for which $J(\{\rho\})$ takes on its minimum value (denoted by I). The problem of designing the PWL switching function can then be considered as a problem in searching the performance surface in order to find that value of $\{\rho\}$ for which $J(\{\rho\}) = I$.

For this work, the cost will be defined as

$$J(\{\rho\}) = \frac{1}{K} \sum_{k=1}^{K} w_k \, T_k(\{\rho\}, \underline{e}_k^o) \qquad (2.2)$$

where k is the index corresponding to the different initial conditions,

K is the number of initial conditions used,

w_k is the weighting coefficient for the k^{th} initial condition,

and T_k is the settling time from $\underline{e}_k^o$ and is dependent upon the switching function parameters $\{\rho\}$.

When the values of w_k and $\underline{e}_k^o$ are fixed by the designer and the summation is carried out, the cost J becomes a function only of the switching function parameters $\{\rho\}$.

The simulation results presented in this investigation were obtained on a hybrid system composed of an IBM 1620 digital computer with 20,000

decimal digits of memory, a 48 amplifier Electronic Associates TR-48 analog computer,and six channels each of A-D and D-A conversion equipment. With careful programming it was possible to run approximately two trajectories per second, including setting new initial conditions and switching function parameter values, sampling the analog voltages to determine when the states had reached the desired region around the origin, and calculating the cost according to Eq. (2.2).

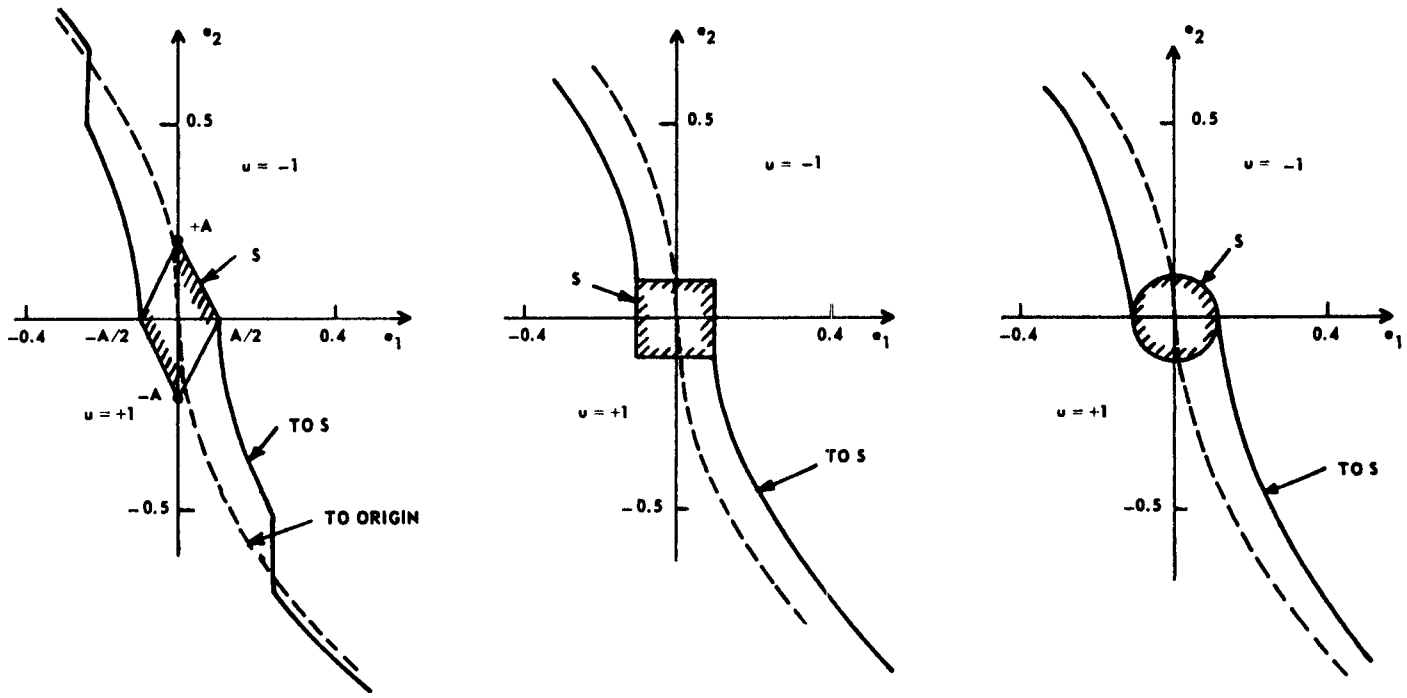
For computational purposes, it is necessary to consider the definition of when a trajectory has reached the state-space origin. Since it is not feasible in an engineering sense to reduce the states to precisely zero, and since the final motion of the trajectory will involve chatter in which the states decay exponentially and hence never precisely reach the origin, it is necessary to establish some cost-free region (denoted by S) surrounding the origin. The resulting performance surface will represent the time required to bring the state to the region S and not to the origin as originally planned.

For second-order plants where the optimal switching curve for trajectories going exactly to the origin can be easily sketched, the optimal switching curve for trajectories going to a region S can be obtained by applying the transversality conditions* to the adjoint vector at the final time, i.e., that time at which the state reaches S. These curves are shown in Fig. 2 for the $1/s^2$ plant, using three representative shapes for the region S having equal areas.

Since virtually the entire theory of minimum-time control systems is for trajectories going to the origin, it is desirable to choose the shape and size of S so that the optimal switching surface corresponding to S will differ as little as possible from that for trajectories going precisely to the origin. With this in mind, the parallelogram

---

* See Sec. 6 of Pontryagin et al. [Ref. 1] for an explanation and derivation. Also see Hutchinson [Ref. 11] for a number of examples.



a. Parallelogram b. Square c. Circle

FIG. 2. OPTIMAL SWITCHING CURVES FOR THE $1/s^2$ PLANT AND THREE CHOICES OF S

shown in Fig. 2a has been used for the simulation of second-order plants. For $n=3$ the region S is a diamond and its definition is readily extendable to higher dimensions.

This choice has two other factors which make it desirable. First, having linear boundaries, it is easily implemented. Second, it is easy to obtain a very close approximation to the difference in optimal times for trajectories going to the origin and to the region S, even for high-order plants. Consider the $1/s^2$ plant for which the optimal switching curves are shown in Fig. 2a. As long as the switching point occurs outside the discontinuities in the optimal switching line, the optimal trajectories going to S will enter it at the points $\underline{e} = \pm\ (0,A)$. Optimal trajectories going directly to the origin will travel along the dashed lines in Fig. 2a which are segments of parabolas and will pass through the points $\underline{e}^T = \pm\ (-A^2/2,A)$ which lie close to the surface of S. Since $|\dot{e}_2| = 1$ for the $1/s^2$ plant, the time required for the state to go from either of these points to the origin on an optimal trajectory is exactly A seconds. Therefore, the minimum time in which the state can reach the points $\underline{e}^T = \pm\ (-A^2/2,\ A)$ is $(T_o - A)$ where T_o is the minimum time to the origin. Since the points $\underline{e}^T = \pm\ (-A^2/2,A)$ differ from the points $\underline{e}^T = \pm\ (0,A)$ by the distance $A^2/2$, when $A \ll 1$ the optimal time to reach S will be approximately $(T_o - A)$ seconds.

For a general second-order plant the above arguments hold when $A \ll 1$ because <u>near the origin</u> the trajectories are very close to those of the $1/s^2$ plant. In addition, the arguments can be generalized to include n^{th}-order plants by using the plant $1/s^n$ in place of $1/s^2$. Therefore, in the simulation work to follow, the optimal times to reach S will be assumed to be given by

$$(T_s)_k = (T_o)_k - A \qquad A \ll 1 \qquad (2.3)$$

where $(T_o)_k$ is the optimal time to reach the origin from the k^{th} initial condition.

In the performance surfaces to be investigated the weighting function w_k will be chosen so that any initial condition will make an equal contribution to the aggregate cost J. This will be done by making $w_k = 1/(T_s)_k$. Therefore, in the results to follow, the cost will be defined as

$$J(\{\rho\}) = \frac{1}{K} \sum_{k=1}^{K} \frac{1}{(T_s)_k} T_k(\{\rho\}, \underline{e}_k^o) \tag{2.4}$$

and $J = 1.0$ will imply that each of the K trajectories is optimal.

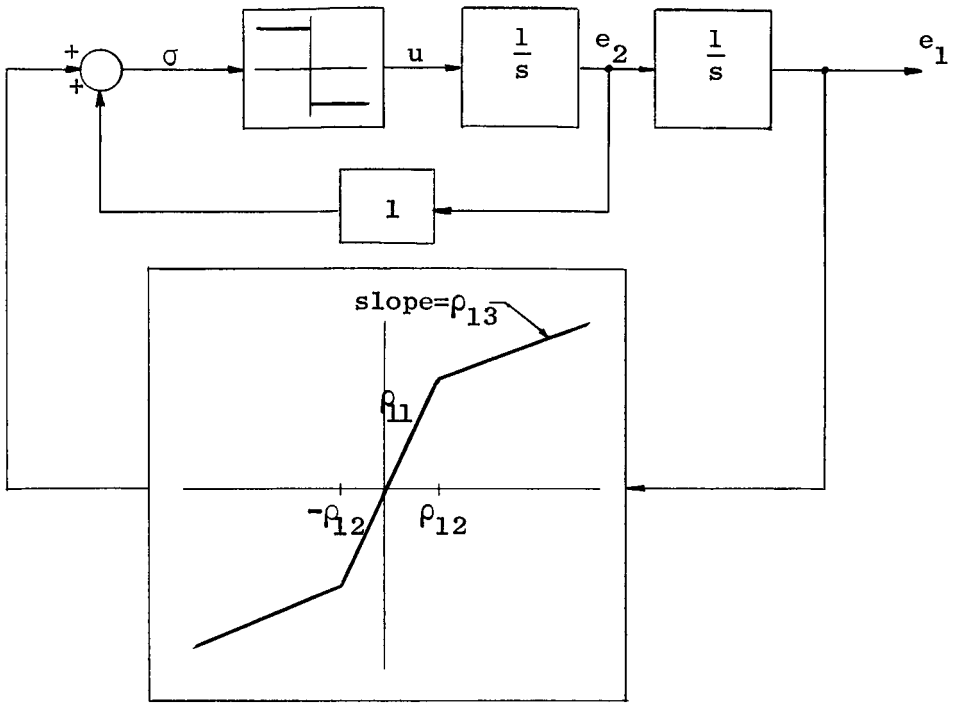
C. PERFORMANCE SURFACES FOR SECOND-ORDER PLANTS

To illustrate some of the rather general statements in the previous section regarding the specification of the performance criterion and its effect upon the surface searching process, the $1/s^2$ plant will be considered. A switching function having $\sigma_1(e_1)$ PWL with one breakpoint, will be used. Thus r=3 and the PWL switching function parameter array can be written as

$$\{\rho\} = \begin{Bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ 1 & x & x \end{Bmatrix} \tag{2.5}$$

where the x's denote undefined elements. The performance surface can be depicted in the four-dimensional space (r+1 = 4) by plotting contour lines of constant J in the plane ρ_{12} = constant. Since ρ_{11} and ρ_{13} are the slopes of σ_1 before and after the breakpoint, it is apparent that the line $\rho_{11} = \rho_{13}$ corresponds to a linear switching function. The block-diagram of the system is shown in the sketch below.

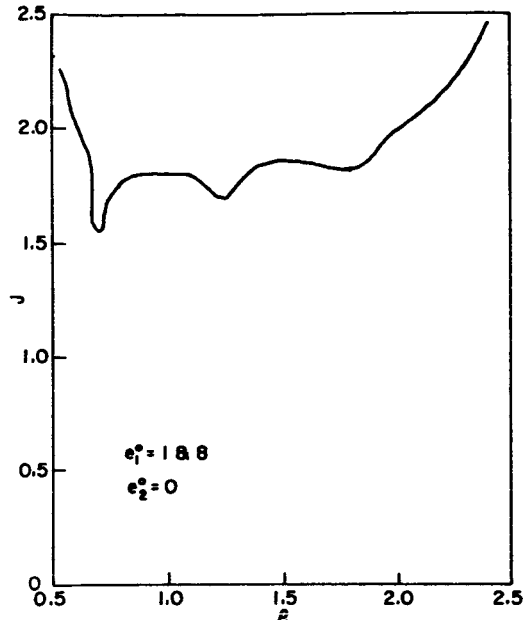
Because the form of $\{\rho\}$ has been fixed, the cost $J(\{\rho\})$ as given by Eq. (2.4) will be completely specified by the choice of the initial conditions to be used, namely the K values for $\underline{e}_k^o$. The choice of



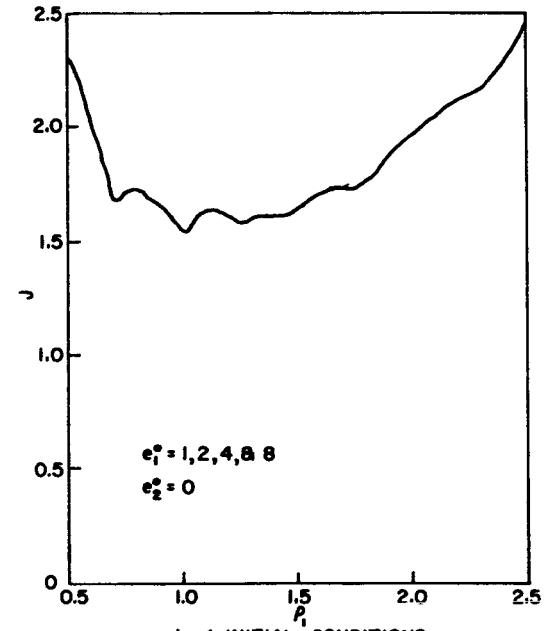
these initial conditions determines the shape of the performance surface which must be searched if the best PWL switching function in the class represented by Eq. (2.5) is to be found. The cost-free region S will be as shown in Fig. 2a, with A = 0.20.

For this particular plant it is possible to limit the initial conditions to points along the positive e_1 axis with no loss in generality because any stable trajectory must intersect the e_1 axis, although this is not true for $n > 2$ and for some other second-order plants. Therefore, for the sake of argument, consideration will be limited to initial conditions along the e_1 axis in the range $0 < e_1^o \leq 8$. Cross-sections of three typical surfaces which were obtained by measuring J at increments of 0.025 in ρ_1 are shown in Fig. 3.

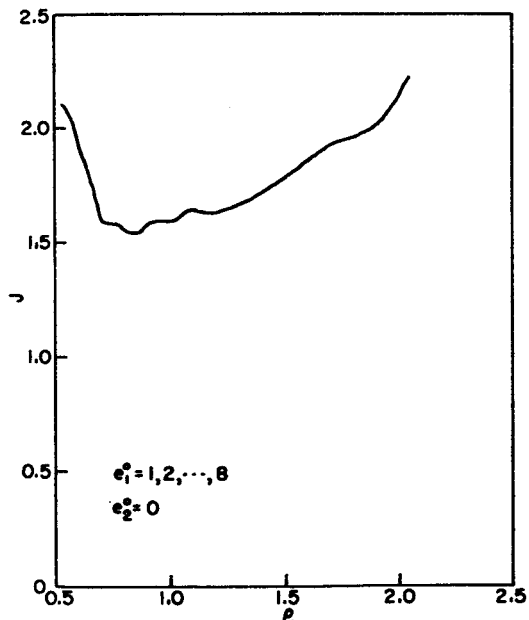
It is found that the surface becomes smoother as the number of initial conditions is increased. The smoothing effect of using more initial conditions in the definition of the cost is explained by noting that the value of J is the average of the cost measured for each of



a. 2 INITIAL CONDITIONS

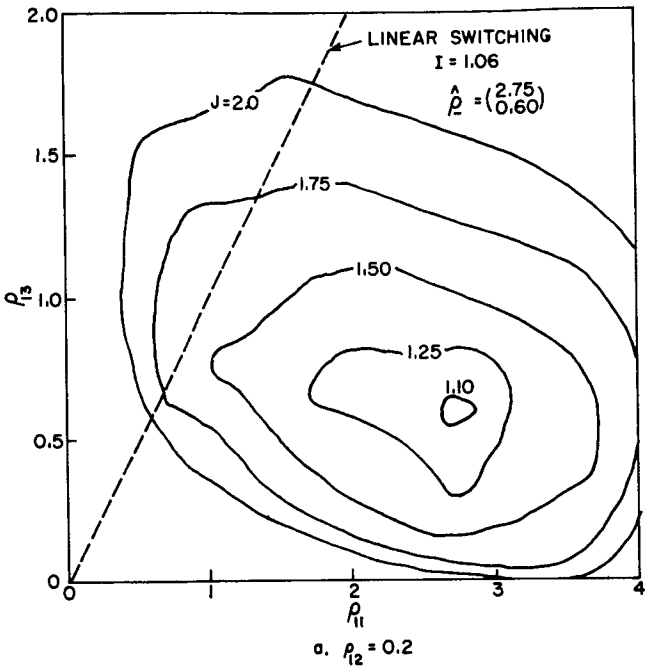


b. 4 INITIAL CONDITIONS



c. 8 INITIAL CONDITIONS

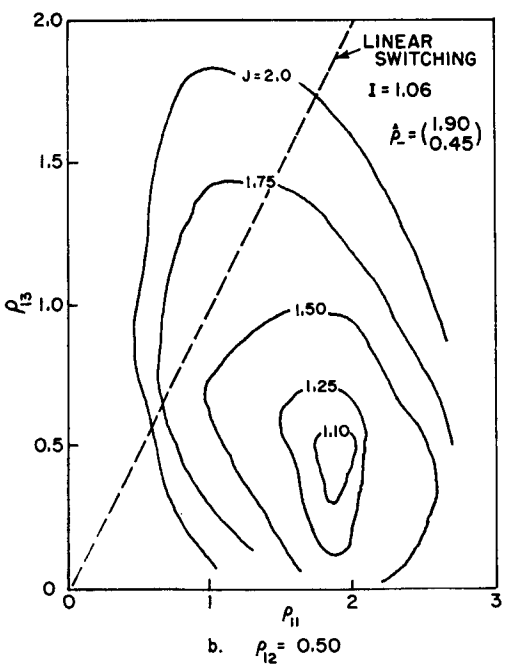
FIG. 3. J VS. ρ_1 FOR 2, 4, AND 8 INITIAL CONDITIONS WITH LINEAR SWITCHING, $1/s^2$ PLANT



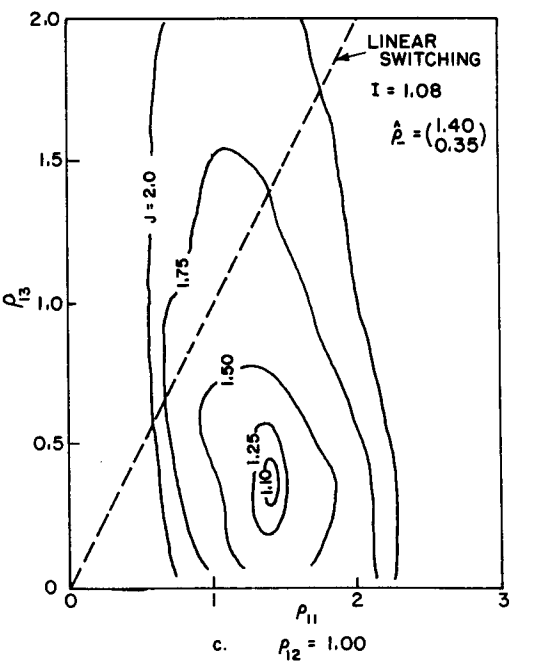
a. $\rho_{12} = 0.2$

FIG. 4. CONTOURS OF J VS. ρ_{11} AND ρ_{13} FOR ρ_{12}=0.20, 0.50 AND 1.0 WITH 8 INITIAL CONDITIONS, $1/s^2$ PLANT

$e_1^o = 1,2,\cdots,8$
$e_2^o = 0$



b. $\rho_{12} = 0.50$



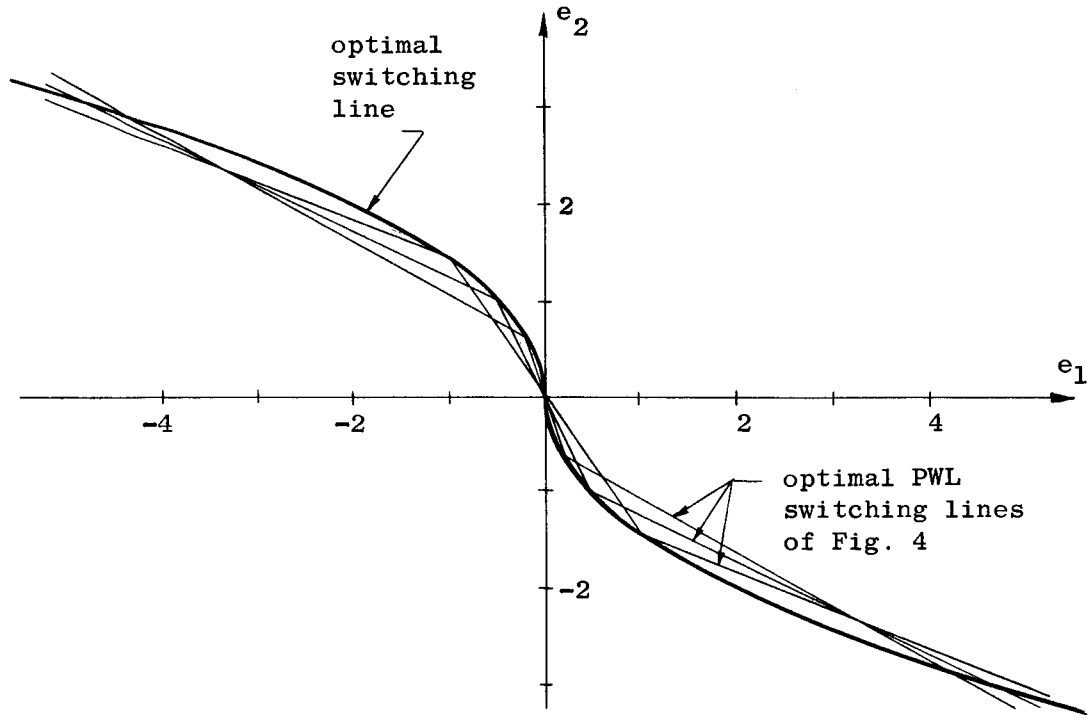
c. $\rho_{12} = 1.00$

the initial conditions used. When $\{\rho\}$ is such that the trajectory from one of the initial conditions is optimal, that particular component of J will be at a minimum. Therefore, there will be a tendency for J to have fluctuations in its gradient and possibly minima at those places in the parameter space where one or more of its components attain their minima. Clearly, the number of places in the $\{\rho\}$ space where this will occur will increase as K, the number of initial conditions used in measuring J, is increased. However, since J is averaged over the K initial conditions, the net effect of any single component upon J will be reduced as K is increased.

Although examination of Fig. 3 indicates that the optimal linear switching parameter depends upon the number of initial conditions used to define J, it was found that the optimal PWL switching parameters did not vary as K was changed from 2 to 8. This indicates that $e_1^o = 1$ and 8 are representative of the entire range $1 \leq e_1^o \leq 8$ and that the only benefit obtained by using more than these two initial conditions is in smoothing the performance surface.

For the curves of Fig. 3 the effect of the number of initial conditions used in determining J has been considered by keeping the breakpoint of σ_1 fixed at $\rho_{12} = 0.20$. It is also of interest to investigate the effect of the breakpoint upon the performance surface. The shape of the performance surface obtained with $K = 8$ is indicated by the contours of J in the ρ_{11}, ρ_{13} plane for three values of ρ_{12} presented in Fig. 4. The point of major interest is that a cost of $J < 1.10$ can be found on <u>each</u> of the three plots, although the values of ρ_{11} and ρ_{13} corresponding to these optimum values differ considerably. Therefore, the three-dimensional parameter space $\{\rho\}$ contains a region within which $J < 1.10$ and this region extends at least from the plane $\rho_{12} = 0.20$ to the plane $\rho_{12} = 1.00$. Since $J \geq 1.00$ by definition and $I \leq 1.08$ for each of the three planes shown, it is apparent that no substantial reduction in cost can be obtained by varying the breakpoint ρ_{12} as long as it lies within the range $0.20 \leq \rho_{12} \leq 1.0$ and the slopes ρ_{11} and ρ_{13} have been set to the optimal values corresponding to that value of ρ_{12} used. However, from an engineering

point of view, the PWL switching functions with $\rho_{12} = 0.20$ or 0.50 are preferable because the minima are broader than when $\rho_{12} = 1.00$. In the sketch below the three PWL switching lines are compared with the optimal switching line which is composed of two portions of parabolas.



In this section it has been shown that the design of PWL switching functions can be treated as a problem in searching a performance surface whose shape will depend upon the choice of initial conditions used in its definition. While it is not practical to give rigorous conditions which are sufficient to ensure that the performance surface can be searched, it has been possible, by considering second-order examples, to formulate the following two rules which would appear to be necessary for the existence of a performance surface which can be searched by one of the standard surface searching techniques. The sufficiency of these rules is demonstrated in specific cases by performing the surface search in order to design linear and PWL switching surfaces for several plants of second-, third-,and fourth-order.

First, a sufficient number of initial conditions should be used so that the size of any local minima and the fluctuations in surface gradient are small enough to be tolerated by the particular searching procedure being used. Second, it is necessary that the initial conditions used in defining J be representative of all of the initial states to which the system might be subjected.

D. SURFACE SEARCHING TECHNIQUES

The task of searching surfaces of the type and dimensionality encountered in designing PWL switching functions by the methods proposed here is far from trivial and is an area of current investigation. Factors which tend to complicate the procedure are the possible presence of relative minima, a complicated functional dependence between J and the PWL switching parameters $\{\rho\}$, and a high dimensionality of the parameter space. A detailed investigation of surface searching techniques has not been undertaken. Instead, two relatively unsophisticated techniques were used to illustrate the feasibility of the search process for representative second- through fourth-order plants with PWL switching. For a detailed discussion of the application of gradient and relaxation procedures to the searching of multiparameter surfaces in the presence of noise, the reader is referred to the work of Kushner [Ref. 12] . Brown [Ref. 13] also gives an excellent discussion of the various gradient procedures for searching surfaces in the absence of noise.

The first of the two search techniques used is a modification of the method of steepest descents. In the method of steepest descents the iterative process for adjusting an r component parameter vector $\underline{R}$ is [Ref. 13]

$$\underline{R}^{m+1} = \underline{R}^{m} - \gamma^{m} (\partial J/\partial \underline{R})^{T} \tag{2.6}$$

where $\underline{R}^{m}$ is the parameter vector before the m^{th} iteration, γ^{m} is the value of a positive coefficient during the m^{th} iteration, and

$\partial J/\partial \underline{R} = (\partial J/\partial R_1, \ldots, \partial J/\partial R_r)$. Equation (2.6) can be considered as representing a discrete feedback process where the coefficient γ is the gain in the feedback loop. If this gain is too low, the convergence of $\underline{R}$ to its optimum value will be slow, and if γ is too high the feedback process will be unstable. Because of the relatively complicated shape of the surfaces to be searched, it proved to be beneficial to make γ dependent upon the results of the process, i.e., to make it be an adaptive parameter. The details of the gradient procedure used in this investigation are presented in Appendix B.

The gradient search procedure worked well when the parameter space was limited to two dimensions, such as the surface depicted in Fig. 4a and in designing a linear switching function for a third-order plant (see Chapter IV). For the performance surface shown in Fig. 4a the search procedure typically reduced the cost from 2.0 to below 1.10 in five iterations. However, when the same procedure was used to search three parameters simultaneously, its convergence was generally quite slow. Due to the geometrical difficulties of analyzing a three-parameter search and the fact that the memory of the available digital computer was virtually saturated, the gradient process was not pursued further. However, it was felt that the simple logic which was successful for two parameters was not sophisticated enough for varying three parameters simultaneously. If more computer memory were available, the methods proposed by Kushner [Ref. 12] would appear to be worth consideration.

When the dimensionality of the parameter space is large, an alternative to using the gradient procedure is a random perturbation search procedure. Because the random search procedure involves perturbations about a nominal point which are small compared to the dimensions of the parameter space, it is distinct from a Monte Carlo approach in which points throughout the parameter space are chosen randomly, in the hope that one will be chosen at or near the optimum value of $\underline{R}$. The random search is particularly attractive for two reasons. First, the logic of the procedure is extremely simple and requires very little computer memory. And second, if $\underline{R}^m$ is caught in a relative minimum of J, there is the possibility that the random perturbation of the parameter

vector will be large enough to move $\underline{R}^{m+1}$ out of the depression surrounding the relative minimum so that it can reach the absolute minimum.

The perturbations of the parameter vector are obtained by adding to $\underline{R}^{m}$ a vector $\underline{\Delta R}_i^m$, each of whose components is an uncorrelated, approximately-Gaussian random variable with zero mean and a standard deviation of $V\sqrt{2}$. The approximately-Gaussian numbers are obtained by adding six random numbers uniformly distributed between -0.5 and 0.5 and multiplying their sum by the positive scale factor V to yield the desired standard deviation [Ref. 14].

The search logic consists simply of measuring the cost corresponding to the perturbed parameter vector and comparing it to the lowest value found by the preceding iterations. A new nominal vector $\underline{R}^{m+1}$ is chosen according to

$$\underline{R}^{m+1} = \underline{R}^{m} + \underline{\Delta R}_i^m$$

if

$$J(\underline{R}^{m} + \underline{\Delta R}_i^m) < J(\underline{R}^{m}) ,$$

otherwise a new random perturbation $\underline{\Delta R}_{i+1}^m$ is chosen. If a value for $\underline{R}^{m+1}$ is not found by the time i reaches the preset limit, N, it is assumed that $\underline{R}^{m}$ is near a minimum and the standard deviation $V\sqrt{2}$ should be reduced. For example, in the method used, after i reaches 32, V is halved and the process repeated, starting again with $i = 1$. If V is halved several times without finding a value of J lower than $J(\underline{R}^{m})$ or reaches some preset level, it is assumed that a minimum has been found. As with the choice of γ for the gradient search, the initial value of V is not critical, due to its adaptive nature, provided that it is not chosen too low.

III. QUALITATIVE DESIGN PROCEDURES FOR LINEAR AND PWL SWITCHING FUNCTIONS

A. GENERAL

When the plant is of third-or higher-order, the number of switching function parameters can lead to a performance surface of such a high dimensionality that it can be an extremely difficult task to search for the optimum. In order to minimize the dimensionality of the performance surface it is very desirable to know which of the linear switching function parameters have the smallest influence upon system performance so that they may be omitted as PWL functions. Another difficulty occurs when the plant has poles which are not well damped, which represents a large portion of the interesting control system problems. In this case the response for values of $\{\rho\}$ in all but a relatively small portion of the r-dimensional parameter space may lead to unstable* trajectories. Since the performance surface does not exist in these regions any search technique dependent upon the surface gradient cannot be applied there. Therefore, attention will be given to methods of obtaining some idea as to where the performance surface exists and to a means of establishing those parameters to which the cost is most sensitive. As an illustration of the procedures developed, second-order examples will be worked before proceeding to the more complicated third- and fourth-order plants for which the methods are intended. At that time several problems peculiar to plants of order greater than two will be discussed.

B. DESIGN OF LINEAR SWITCHING FUNCTIONS

The method of designing an optimal** PWL switching function will be to start the search procedure with a linear switching function which

---

*Unstable will be used to imply that the state-space origin is not asymptotically stable.

**An optimal PWL switching function is the optimal function for the PWL components σ_i and number of breakpoints corresponding to $\{\rho\}$.

works reasonably well for all of the K initial conditions and then to modify it by making one or more of its components PWL. However, there is no general procedure for the synthesis of linear switching functions such as there are for strictly linear systems.

As was mentioned in Section I-B, Flügge-Lotz and her co-workers [Refs. 7, 9] have synthesized linear switching functions for particular third-order plants and particular types of initial conditions by choosing the linear switching coefficients ρ_i to be a good approximation to those coefficients which reduced the state to the origin in (n-1) switchings as one component of the initial state was varied over the range of interest. While this approach can give good results in specific instances it has several drawbacks which limit its applicability. First, for plants with oscillatory roots there are large regions of stable response from which the origin cannot be reached in (n-1) switchings. Second, the value of $\underline{\rho}$ corresponding to a particular initial condition must be determined, presumably by computation of trajectories or computer simulation. Third, the effect of using a linear switching function designed on the basis of initial states along the e_1 axis for slightly different initial states can result in unstable trajectories when the plant is of third-or higher-order. Finally, the manner in which the variable coefficients obtained by considering different initial states are to be approximated by the constants ρ_i and the consequences of this approximation are not apparent without further simulation of the system.

Schmidt [Ref. 3] has used the root-locus method to study saturating linear systems of third-and higher-order by treating the limiter as an equivalent gain which decreases after it reaches saturation. This interpretation yields an equivalent set of closed-loop poles which coincide with the closed-loop poles of the unsaturated linear system when the limiter is not in saturation and travel along the root loci toward the open-loop plant poles as the limiter goes further into saturation. He has shown by simulation studies that the root loci can be a useful tool for analyzing the response of this type of system in a qualitative manner. Following this approach will allow a qualitative interpretation

of the switching plane in terms of the root loci and permit one to obtain useful information regarding both linear and PWL switching functions.

Kalman [Ref. 15] has investigated the stability of quite arbitrary nonlinear systems by applying the root-locus techniques to the incremental differential equation. However, when the nonlinearity is a contactor the only two incremental equivalent gains are zero and infinity. For this reason, it appears to be more reasonable to choose the equivalent gain so that its output coincides with that of the contactor for the same value of σ. Since $u = -\text{sgn}\ \sigma$ and $\text{sgn}\ \sigma = -\sigma/|\sigma|$, the desired equivalent gain is $K(\sigma) = -1/|\sigma|$. For the sake of comparison, the describing function for the contactor is $-4/\pi|\sigma|$ which is approximately the same as $-1/|\sigma|$. Since no quantitative results are to be obtained by using this equivalent linearization the small difference between the two is inconsequential.

The linear switching function can be expressed as

$$\sigma(\underline{e}) = \rho_1 e_1 + \rho_2 e_2 + \ldots + \rho_n e_n \ . \tag{3.1}$$

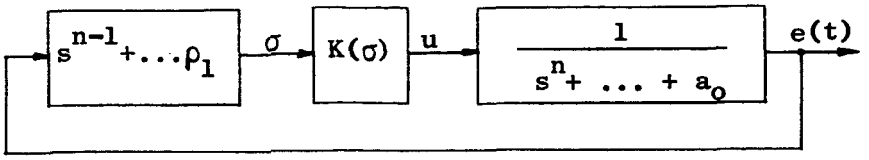
Since the switching function is followed by the contactor, one of the ρ_i may have any positive value without affecting $u(t)$, so ρ_n will be set equal to unity in the work to follow. If the vector $\underline{e}$ is composed of the error variable and its first $(n-1)$ derivatives, as is often the case, then

$$\sigma(\underline{e}) = \frac{d^{n-1}e}{dt^{n-1}} + \ldots + \rho_2 \frac{de}{dt} + \rho_1 e \ . \tag{3.2}$$

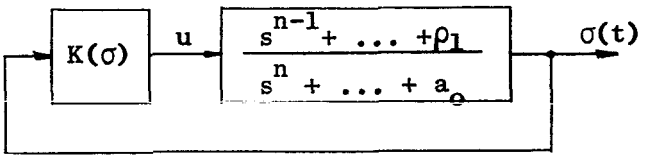
Taking the Laplace transform of Eq. (3.2) yields the transfer function

$$\frac{\Sigma(s)}{E(s)} = s^{n-1} + \ldots + \rho_2 s + \rho_1 \tag{3.3}$$

and the total system can be represented by the block-diagram shown in Fig. 5a. The fact that the transfer function (3.3) obviously cannot



a. e(t) as Output Variable



b. σ(t) as Output Variable

FIG. 5. BLOCK-DIAGRAMS OF THE LINEARIZED SYSTEM

represent a physical system, since it has no poles, is of no concern because it has been assumed that the (n-1) derivatives of e have been generated by the plant and the switching function represents only their linear combination.

The zeros of the transfer function can be found by factoring the right-hand side of Eq. (3.3) into the form

$$s^{n-1} + \ldots + \rho_1 = \begin{cases} (s-\alpha) \prod_{i=1}^{(n-2)/2} (s^2 + 2\zeta_i \omega_i s + \omega_i^2) & n \text{ even} \\ \prod_{i=1}^{(n-1)/2} (s^2 + 2\zeta_i \omega_i s + \omega_i^2) & n \text{ odd} \end{cases} \tag{3.4}$$

where the ω_i and ζ_i are the natural frequencies and damping ratios, respectively, of each pair of zeros and, if n is even, α describes a single real zero.

When $n > 3$ in Eq. (3.3) or when the system is composed of two or more coupled plants, it will be more convenient to use the parameter vector $\underline{\xi}$ which locates the zeros of the transfer function from the control u to the switching function σ, rather than $\underline{\rho}$, the linear switching function coefficients. The $(n-1)$ components of $\underline{\xi}$ are defined as

$$\xi_j = \begin{cases} \omega^2_{(j+1)/2} & j = 1,3,\ldots,n-2 \\ 2\zeta_{j/2}\,\omega_{j/2} & j = 2,4,\ldots,n-1 \\ \alpha & j = n-1 \text{ if } n \text{ even.} \end{cases} \tag{3.5}$$

If the blocks denoting the switching function and plant in Fig. 5a are combined, so that the error is no longer considered a variable, the pseudo-system of Fig. 5b is obtained, having $(n-1)$ zeros, n poles, and the variable gain $K(\sigma) = -1/|\sigma|$. This system, whose output is the scalar function $\sigma(t)$, can be analyzed in a qualitative manner by considering the root loci. The entire range of the equivalent gain $K(\sigma)$ from zero to infinity is of interest.

Since σ is the linear combination of the n components of $\underline{e}$ given by Eq. (3.1), $\underline{e} = 0$ implies that $\sigma = 0$ but the converse is not true. This being the case, there are two conditions which are necessary in order for the state-space origin to be asymptotically stable. First, the average values of σ and $\dot{\sigma}$, the output of the switching function and its derivative, must both be reduced to zero. Second, the chatter motion when $\sigma = 0$ must be such that $\underline{e}$ approaches zero.

It is well known that the behavior of $e(t)$ during chatter can be easily determined for systems of any order by setting σ in Eq. (3.2) equal to zero, yielding the following $(n-1)^{st}$-order homogenous differential equation for e:

$$\frac{d^{n-1}e}{dt^{n-1}} + \ldots + \rho_2 \frac{de}{dt} + \rho_1 e = 0$$

Examination of this differential equation indicates that the average motion in chatter is precisely that of a linear system having closed-loop poles at the zeros corresponding to the switching function. For the chatter motion of $\underline{e}$ to be asymptotically stable the switching function zeros must be located in the left half of the s-plane (hereafter abbreviated LHP), implying that the coefficients of s in Eq. (3.4) must satisfy the Routh-Hurwitz criterion.

The first condition, namely, that σ and $\dot{\sigma}$ be reduced to zero, falls in the realm of the celebrated Aizerman's conjecture which deals with the stability of systems of the type shown in Fig. 5b. As described by Hahn [Ref. 16] , the conjecture states that if a single nonlinearity $f(\sigma)$ is replaced by a linear gain and the linearized system is asymptotically stable for all values of the gain between the limits K' and K'', then the nonlinear system is asymptotically-stable-in-the-large for any continuous single-valued nonlinearity satisfying the condition

$$K' < \frac{f(\sigma)}{\sigma} < K'' \quad .$$

Hahn states that the conjecture has been proven valid for all second-order plants but invalid for certain third-order plants having two zeros in their transfer functions.

The version of Aizerman's conjecture restricted to a contactor nonlinearity will be used to establish design criteria on the root loci and those regions of the s-plane in which they may lie. If the root loci remain entirely in the LHP for all $0 < K(\sigma) < \infty$ it will be anticipated that the nonlinear system is asymptotically-stable-in-the-large. If the loci enter the RHP as the equivalent contactor gain decreases and remain there for all lower values of gain, it will be anticipated that the nonlinear system will yield unstable responses for some initial conditions in the state-space. Should the loci correspond to those of a conditionally stable system it will be anticipated that the system has a finite region of stability, although Kalman [Ref. 15] states that if the loci enter the RHP only slighly and then leave,

there may be no actual unstable region in the state-space. Generally speaking, the criterion is followed that the equivalent poles, which are functions of σ, should remain as far to the left in the s-plane as possible.

This general rule can be made more specific when the plant has roots near the imaginary axis. In that case, a function which has considerable utility is the angle with which the locus departs from an oscillatory pole. This angle, denoted by ϕ, is measured from a line parallel to the real axis of the s-plane and passing through the oscillatory pole having positive imaginary part, as indicated in Fig. 6. If

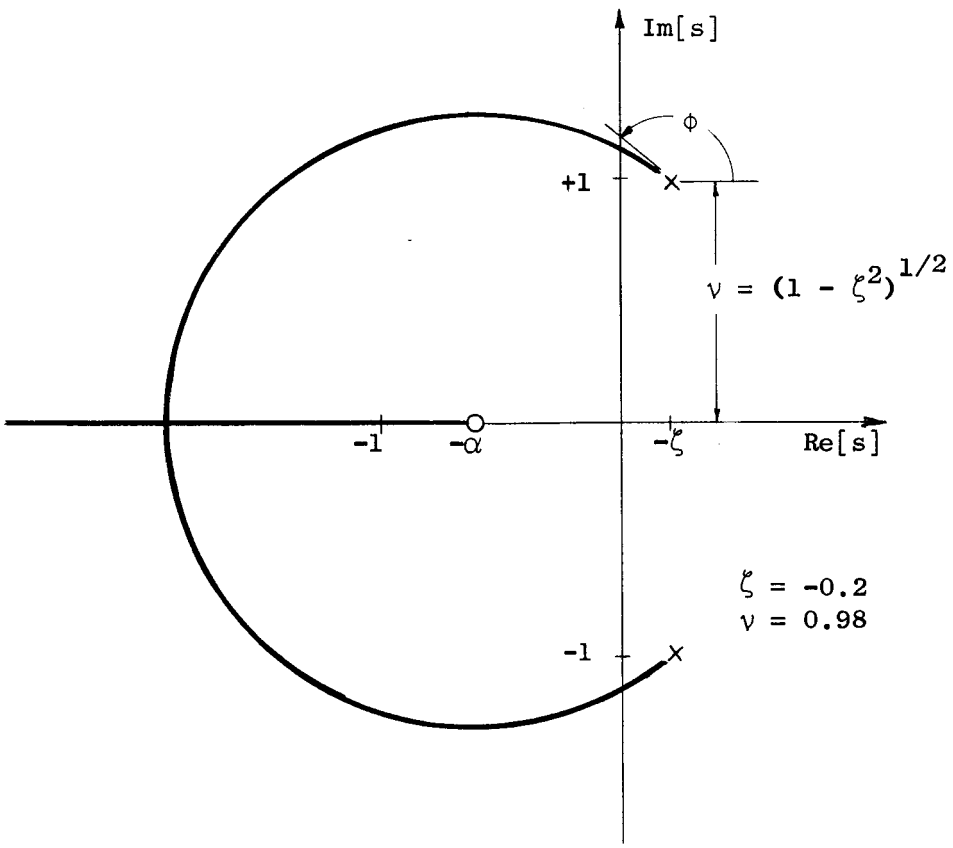


FIG. 6. TYPICAL ROOT-LOCUS PLOT FOR THE $1/(s^2 - 0.4s + 1)$ PLANT

the oscillatory poles in question lie exactly on the imaginary axis then, unless the switching function parameters place ϕ in the range $90^{o} < |\phi| < 270^{o}$ the linearized system cannot be asymptotically-stable-in-the-large since the loci enter the RHP.

It should be emphasized that Aizerman's conjecture will be used only in a qualitative sense to obtain parameter values for initializing a surface searching procedure. Therefore, any conclusions regarding the stability of a nonlinear system obtained by applying the conjecture would be verified at the start of the search procedure because the measurement of the performance surface requires the computation of the system trajectories.

In addition to the design criteria derived from the root locus, it is possible to obtain one more criterion by considering the trajectories in the n-dimensional state-space. It has been observed during a large number of analog computer runs with a variety of third- and fourth-order plants and linear switching functions that the stability of a nonlinear system is characterized by the existence or absence of a unique symmetric periodic solution having two control reversals each period.* If the plant poles lie in the LHP the periodic solution, henceforth referred to as the dominant periodic solution, is stable and it experimentally is observed that large initial states will converge to it and never reach the origin. When all of the plant poles are not in the LHP it is observed that the dominant periodic solution is unstable and initial states lying outside the stability boundary in the n-dimensional state-space will approach the dominant periodic solution before growing without bound. In either case, it has been observed that the absence of the dominant periodic solution implies that the system will be stable for any initial conditions which can be simulated on the analog computer.

---

*Other symmetric periodic solutions involving more than two reversals may possibly exist, but only those having two reversals will be considered here. These solutions are the only periodic motions consistently observed in all classes of systems being simulated. When n=2 the dominant periodic solution is the limit cycle having two control reversals per period.

If the trajectories of the plant are examined in a canonical space where the projections of the trajectories have a relatively simple geometry, the dominant periodic solution can generally be deduced by using the conditions of symmetry and periodicity in order to obtain a locus of possible switching points in the state-space. Since the switching points must lie on the switching plane the periodic solution can occur only when the locus of possible switching points intersects the switching plane. For this reason it is possible to express some measure of the size of the periodic solution, such as the amplitude of a state variable, in terms of the switching function parameters.

The periodic solutions obtained in this manner are identical to those obtained by construction of the Hamel or Tsypkin* loci. While the Hamel and Tsypkin loci can be applied in a straightforward manner to systems having nonlinearities considerably more general than a perfect contactor, they are not in a convenient form for use with switching surfaces having more than two dimensions.

To summarize, the criteria for the qualitative design of a linear switching function are:

1. Keep the switching function zeros as far to the left in the s-plane as possible by minimizing their maximum real part.

2. Keep the equivalent poles of the linearized system as far to the left in the s-plane as possible by optimizing a suitable measure such as the angle of departure or cross-over gain on the root loci.

3. Maximize the size of the dominant periodic solution and, if possible, eliminate it entirely.

One or more points in that region of the $\underline{\xi}$ space within which <u>all</u> of the three criteria are satisfied in a reasonable manner can be used to initiate the search of the performance surface for the optimal linear or PWL switching function parameters, or the cost can be evaluated at several points to determine a linear switching function which will presumably be close to the optimal.

---

*See Chapter 26 of Gille, Pelegrin, and Decaulne [Ref. 17] .

The methods described in the preceding section for analyzing linear switching functions will be applied to the second-order plant with transfer function $1/(s^2 + 2\zeta s + 1)$ in order to show their usefulness and the interpretation of the results. The differential equation describing the system is

$$\dot{\underline{e}} = \begin{pmatrix} 0 & 1 \\ -1 & 2\zeta \end{pmatrix} \underline{e} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u \qquad (3.6)$$

where $\underline{e}^T = (e, \dot{e})$ and $|u| = 1$.

It is easier to construct the phase-plane trajectories of the system if the canonical variables defined by the following transformations are used:*

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & \zeta \\ 0 & \nu \end{pmatrix} \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} \quad ; \quad \begin{pmatrix} e_1 \\ e_2 \end{pmatrix} = \begin{pmatrix} 1 & -\zeta/\nu \\ 0 & 1/\nu \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad (3.7a) \quad (3.7b)$$

where $\nu = \sqrt{1 - \zeta^2}$ and $|\zeta| < 1$.

The trajectories of the canonical variables are logrithmic spirals with foci at $\underline{x}^T = \pm (u, 0)$.

Since the plant is of second-order the linear switching function given by Eq. (3.1) takes the form

$$\sigma = e_2 + \alpha e_1 \qquad (3.8)$$

and Eq. (3.3) becomes

$$\frac{\Sigma(s)}{E(s)} = s + \alpha . \qquad (3.9)$$

---

*See pp. 124-129 of Flügge-Lotz and Yin [Ref. 18] for a derivation.

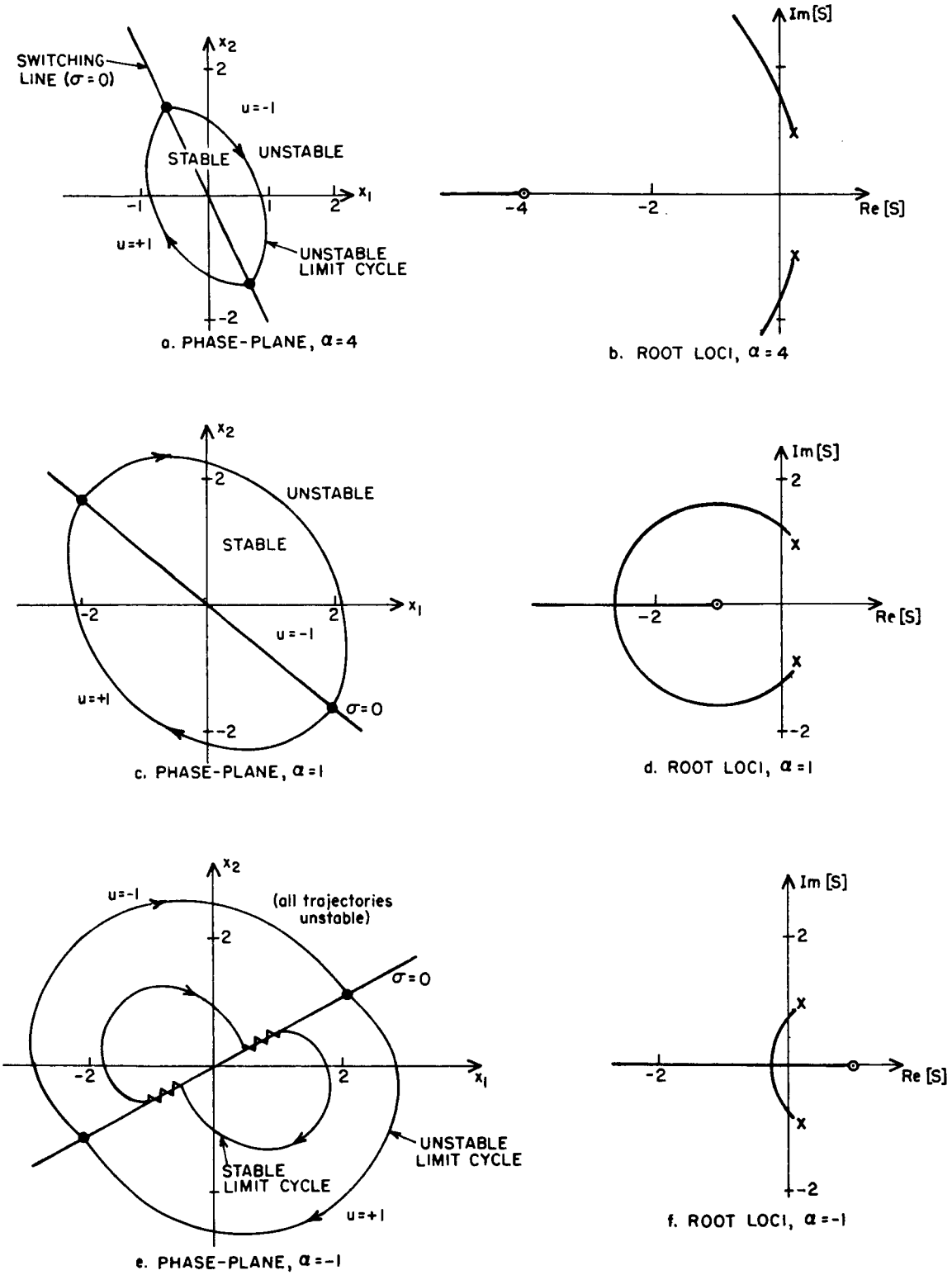


FIG. 7. PHASE-PLANE AND ROOT-LOCUS PLOTS FOR THE $1/(s^2-0.4s+1)$ PLANT WITH LINEAR SWITCHING

Therefore, the switching function is represented in the s-plane by a single real zero and the plant by two poles whose location depends upon the parameter ζ.

Consider the situation when $\zeta = -0.2$. As seen in Fig. 6, the poles are a complex pair slightly into the RHP so that some portion of the root loci must lie in the RHP, no matter what value α has. According to the arguments of the previous section there will be some region of the phase-plane for which the trajectories are unstable for any linear switching function. In Fig. 7 the stability boundaries in the canonical phase-plane and the corresponding root-locus plots are shown for several representative values of α. By setting Eq. (3.8) equal to zero and applying the transformation of Eq. (3.7b), the equation of the switching line in the canonical space is found to be

$$x_2 = - \left(\frac{\alpha v}{1 - \zeta} \right) x_1 .$$

The manner in which the equivalent linear system, i.e., the root-locus plot, depicts the salient features of the nonlinear system, i.e., the phase-plane plot, is described below:

1. Figures 7a and b ($\alpha = 4$). The zero is well in the LHP, implying that the chatter motion decays rapidly (time constant = 1/4 second). However, the angle of departure of the locus from the upper pole is only slightly over 90° meaning that the locus exists for an appreciable distance in the RHP. This implies that the stability region may be smaller than that for $\alpha = 1$, where ϕ has a larger value. Comparison of the stability regions in Figs. 7a and c verify that this is the case.

2. Figures 7c and d ($\alpha = 1$). The zero is still well in the LHP and ϕ has a larger value than for $\alpha = 4$, implying that the chatter motion is stable and the region of stability should be larger, which it is.

3. Figures 7e and f ($\alpha = -1$). Again the loci leave the poles with ϕ considerably closer to 180° than when $\alpha = 4$, implying a large stability region. However, the zero in the RHP means that the chatter is unstable and the state can never reach the origin and remain there.

The phase-plane plot indicates that the nonlinear system has gained a stable limit cycle toward which any trajectory starting inside the stability boundary will move, even one starting at the origin.

Of the three possible switching functions shown, the one with $\alpha = 1$ would presumably represent the best choice, a conclusion which can be drawn readily from the root loci. If $0 < \alpha < 1$, the chatter motion is stable but the time constant becomes large, approaching infinity as α approaches zero. However, the point of this discussion has been only to illustrate the application of the design procedure and hopefully to give the reader some confidence in its utility. It will be found to be considerably more useful in the higher-order examples to be worked later.

It is interesting to note that if the sign of the switching function had been reversed (i.e., $\rho_n < 0$) the feedback would be positive and one would expect the nonlinear system to be unstable, as it is. The root loci reflect this situation by requiring use of the 0^o loci rather than the 180^o loci. Since the positive real axis to the right of the last real singularity is always on the 0^o loci, such a system would always be unstable.

C. DESIGN OF PWL SWITCHING FUNCTIONS

From the discussion of the previous section it is apparent that as the zeros corresponding to the linear switching function are moved to the left in the s-plane the rate of decay of the chatter motion is increased and presumably the cost for small disturbances is reduced. However, as they are moved further to the left the zeros exert a decreasing influence upon the loci departing from the poles. If the plant has three or more poles and no zeros are present, at least one of the loci will enter the RHP. Shifting the zeros well to the left of the poles will make the loci approach these "no-zero-loci." Therefore, when $n \geq 3$, there is a limit to how far the zeros may be moved to the left in the s-plane without causing one or more of the trajectories to become unstable.

In an effort to have both a rapidly decaying response for small disturbances and a large region of stability it is reasonable to consider moving the switching function zeros to the right or left in the s-plane depending upon the state of the system.

Linear switching constrains the designer to a fixed set of zeros, but the PWL switching function can be treated in a qualitative sense as a shifting of the zeros, dependent upon the state vector $\underline{e}$. This shifting of the zeros is accounted for by treating the PWL functions $\sigma_i(e_i)$ as equivalent gains which are raised or lowered when e_i exceeds the break points. This same procedure has been used by Schmidt [Ref. 3] to compute nonlinear functions of a single state variable such that the switching function zeros move in a manner which is deemed desirable. Because the shifting of the zeros by means of the PWL functions is a second equivalent linearization beyond the linearization of the contactor, neither one of which can be justified on a rigorous basis, any results derived from its application should be used with caution and must be subjected to computer verification.

The first problem to be solved in the design of a PWL switching surface is the selection of the component of the switching function to be made PWL and of reasonable values which the parameters describing this component should have. These preliminary steps are essential if a performance surface suitable for optimization is to be found. The general criteria described in the previous section will be followed in solving this problem.

If it is possible to express the desired zero shifting in terms of a root-locus index such as angle of departure whose value is to be increased or decreased as the zeros shift, then the partial derivatives of this quantity with respect to the linear switching function parameters may be evaluated. By knowing the partial derivatives and also the approximate ranges over which the state variables will vary, the designer can obtain a qualitative indication of the manner in which the switching function components should be made PWL.

According to the above arguments, there are two conditions which are necessary in order to improve the system performance by making one of the components of σ, say $\sigma_i(e_i)$ PWL. First, the partial derivative of the root-locus index with respect to ρ_i should be as large as possible. Second, there must be a large enough difference in the range over which e_i varies for large and small disturbances to insure that the change(s) in slope at the breakpoint(s) of $\sigma_i(e_i)$ will affect the responses for large disturbances but not those for small disturbances.

Also, the switching function can be made PWL on the basis of the dominant periodic solution. This is accomplished by choosing the PWL parameters so that the intersection of the switching surface and the locus of possible periodic solution switching points yields as large a periodic solution as possible. It is clear that if the periodic solution corresponding to a particular linear switching function is to be enlarged or eliminated, the breakpoint of the PWL σ_i must be smaller than $|e_i|$ at its switching points.

The partial derivatives of the angle of departure of the locus from an oscillatory pole with respect to the switching function can be evaluated in the following manner. The angle of departure is a property of the root locus and, since the pole locations are fixed, it must be a function of the (n-1) zero parameters ξ_i. Furthermore, the ξ_i define the roots of the $(n-1)^{st}$-order polynomial with coefficients ρ_j as indicated in Eq. (3.4). Therefore, the n partial derivatives defined by

$$\frac{\partial \phi}{\partial \underline{\rho}} = \left(\frac{\partial \phi}{\partial \rho_1} , \frac{\partial \phi}{\partial \rho_2} , \ldots , \frac{\partial \phi}{\partial \rho_n} \right)$$

may be found by applying the chain rule of partial differentiation,

$$\frac{\partial \phi}{\partial \rho_j} = \sum_{i=1}^{n-1} \frac{\partial \phi}{\partial \xi_i} \frac{\partial \xi_i}{\partial \rho_j} \qquad j = 1,2,..,n. \tag{3.10}$$

If $\rho_n = 1$ the value of $\partial\phi/\partial\rho_n$ cannot be found by applying Eq. (3.10) because ρ_n will not appear as a variable in the equations for ξ_i. On the other hand, since there is one arbitrary relationship between the ρ_j there must be a corresponding relationship between the components of $\partial\phi/\partial\underline{\rho}$. To derive this relationship it is noted that, since the switching function is followed by a contactor, the linear switching coefficients ρ_j can be multiplied by any positive constant without affecting either $u(t)$ or the locations of the zeros.

Therefore, if each component of $\underline{\rho}$ is multiplied by the constant $(1+\epsilon)$ where ϵ is arbitrary, it follows that

$$\phi(\underline{\rho} + \epsilon\underline{\rho}) = \phi(\underline{\rho}) .$$

If $\epsilon \ll 1$, $\phi(\underline{\rho} = \epsilon\underline{\rho})$ may be expanded about the point $\underline{\rho}$ and all terms of order ϵ^2 and higher dropped, yielding

$$\phi(\underline{\rho} + \epsilon\underline{\rho}) = \phi(\underline{\rho}) + \epsilon \sum_{j=1}^{n} \rho_j \frac{\partial\phi}{\partial\rho_j} .$$

Since the above two equations must hold for arbitrary ϵ, it follows that

$$\sum_{j=1}^{n} \rho_j \frac{\partial\phi}{\partial\rho_j} = 0, \qquad (3.11)$$

so the n^{th} component of $\partial\phi/\partial\underline{\rho}$ can be found if the other $(n-1)$ are known. Using Eqs. (3.10) and (3.11), the equations for the n partial derivatives can be written as

$$\frac{\partial\phi}{\partial\rho_j} = \sum_{i=1}^{n-1} \frac{\partial\phi}{\partial\xi_i} \frac{\partial\xi_i}{\partial\rho_j} \qquad j = 1,2,\ldots,n-1 \qquad (3.12a)$$

and

$$\frac{\partial \phi}{\partial \rho_n} = - \frac{1}{\rho_n} \sum_{j=1}^{n-1} \rho_j \frac{\partial \phi}{\partial \rho_j} . \qquad (3.12b)$$

A simple example of the calculation of $\partial\phi/\partial\underline{\rho}$ and its use in the PWL design is given by the plant with transfer function $1/(s^2 - 0.4\ s + 1)$ which was treated in the previous section. From Eq. (3.5), with $n = 2$, it follows that the zero parameter vector $\underline{\xi}$ is simply the scalar $\xi_1 = \alpha$. By comparing Eqs. (3.3) and (3.9) it is apparent that the linear switching parameter vector $\underline{\rho}$ is given by $\underline{\rho}^T = (\alpha, 1)$. Since the root locus in the vicinity of the poles is a circle centered at the zero (see Fig. 6), the angle of departure is given by

$$\phi = \frac{\pi}{2} + \arctan \frac{\nu}{\xi_1 - \zeta} .$$

Because $n = 2$ and $\xi_1 = \rho_1$, Eq. (3.12a) reduces to

$$\frac{\partial \phi}{\partial \rho_1} = \frac{\partial \phi}{\partial \xi_1}$$

and Eq. (3.12b) reduces to

$$\frac{\partial \phi}{\partial \rho_2} = - \rho_1 \frac{\partial \phi}{\partial \rho_1} .$$

Solving for $\partial\phi/\partial\xi_1$ and substituting α for ξ_1 yields

$$\frac{\partial \phi}{\partial \underline{\rho}} = \left(\frac{-\nu}{\alpha^2 - 2\zeta\alpha + 1} , \frac{\alpha\nu}{\alpha^2 - 2\zeta\alpha + 1} \right) .$$

Figure 8 shows the two components of $\partial\Phi/\partial\underline{\rho}$ plotted against the zero location α, where the values of $\zeta = -0.2$ and $\nu = 0.98$ have been used. The fact that $\partial\Phi/\partial\rho_1 < 0$ when $\alpha > 0$ implies that the effective value of Φ can be increased by making σ_1 be PWL, with a decrease in slope at the breakpoint. Likewise, making σ_2 be PWL so that the slope is increased at its breakpoint should improve the size of the stability region.

The other condition necessary for the system performance to be improved by making a particular switching function component σ_i PWL is that e_i must range over a set of values large enough so that it would exceed the breakpoints of σ_i. For this particular plant both components of $\underline{e}$ will become large if either one does, due to the oscillatory nature of the poles. Therefore, one would expect that making either one of the σ_i PWL would improve the region of stability while retaining a relatively good small-disturbance response. This should be no surprise because, as was mentioned in Sec. II-A, when the state-space is only two-dimensional the same PWL switching function can be obtained by making either one of the components PWL.

To give a qualitative picture of the effect of making σ_1 a PWL function with one breakpoint consider Fig. 9. The linear portion of the switching function corresponds to $\alpha = 4$ which results in a good small-disturbance response but a small region of stability (see Fig. 7a). The breakpoint and change in slope of σ_1 have been chosen so that the effective value of Φ has been raised for the larger states. Examination of the sample trajectories indicates that the origin is stable for all states within the limit cycle shown, which is virtually the largest stability region obtainable with $|u| \leq 1$. Therefore, the effect of the PWL switching function shown has been to provide a substantial improvement in the combined large- and small-disturbance response over that obtainable with linear switching, at only a small increase in switching function complexity.

To illustrate the second design criterion the PWL switching function is examined from the point of view of its effect upon the size of the periodic solution. In this second-order case the periodic solution is

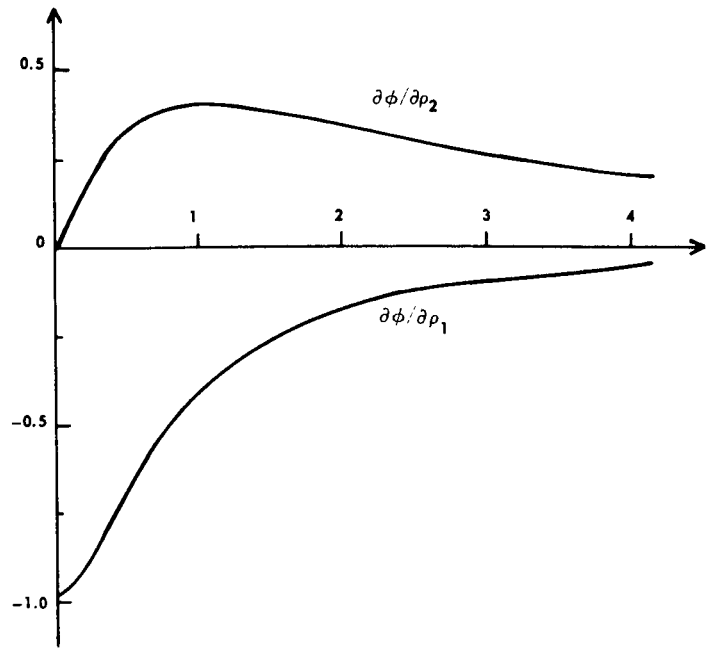


FIG. 8. $\partial\phi/\partial\rho_1$ AND $\partial\phi/\partial\rho_2$ VS. α FOR THE $1/(s^2-0.4s + 1)$ PLANT

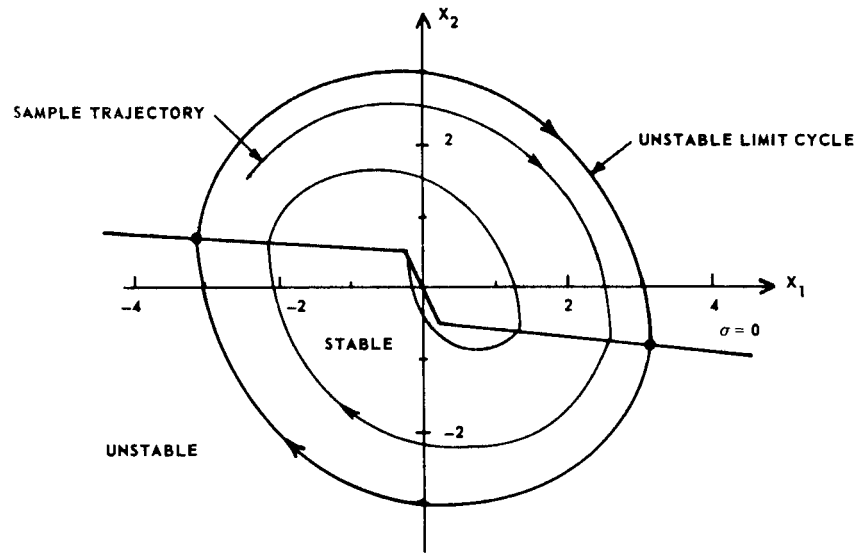


FIG. 9. PHASE-PLANE PLOT FOR THE $1/(s^2-0.4s+1)$ PLANT WITH PIECEWISE-LINEAR SWITCHING

the unstable limit cycle outside of which all trajectories grow without bound. The equations for the locus of possible switching points in the canonical space, expressed in terms of the quarter-period ψ, are*

$$x_1 = \pm \frac{-\sinh 2\zeta\psi}{\cos 2\nu\psi + \cosh 2\zeta\psi} \quad \text{and} \quad x_2 = \mp \frac{\sin 2\nu\psi}{\cos 2\nu\psi + \cosh 2\zeta\psi} .$$

When linear switching is used with $\alpha = 4$ the switching line intersects the locus of possible switching points shown in Fig. 10 at the points N_1 and N_2, and the resulting stable region is as shown in Fig. 7a. When the switching function is made PWL so that $\sigma = 0$ along the dashed line, the intersections of the switching line and the locus of periodic solution switching points move from N_1 and N_2 to N_1' and N_2' giving the considerably larger region of stability shown in Fig. 9.

Finally, the merits of the PWL switching function over linear switching can be evaluated on the basis of how well each one approximates the optimal switching curve. In Fig. 11 the optimal curve, which was first found by Bushaw [Ref. 19], is compared with the linear and PWL switching lines discussed in connection with Fig. 9. It is apparent that the PWL curve yields a considerably better approximation to the optimal curve than does the linear switching line.

---

*The equations in terms of e_1 and e_2 are given by Flügge-Lotz [Ref. 8] and can be transformed to the canonical variables by using Eq. (3.7b).

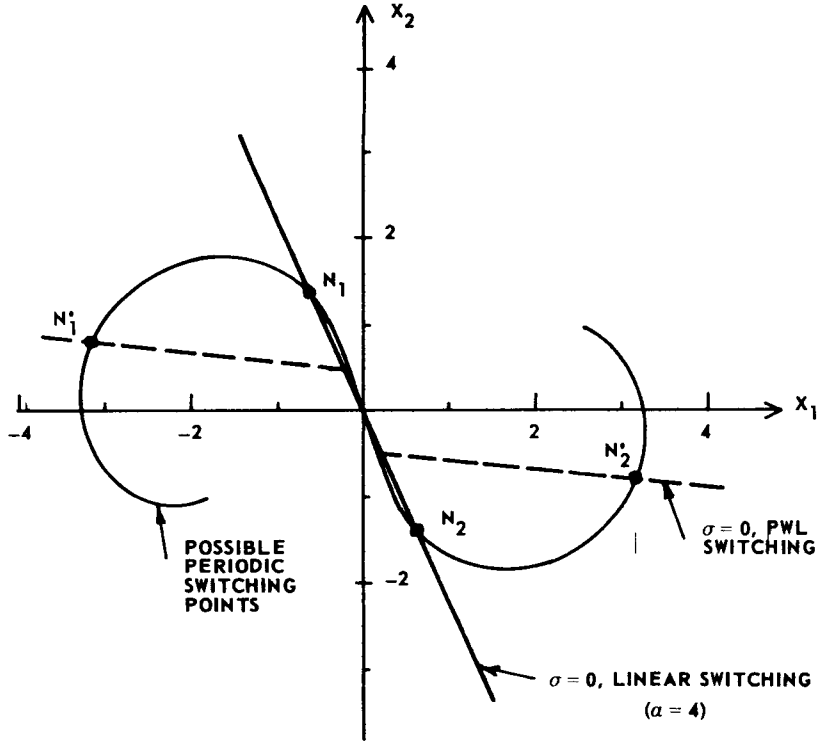


FIG. 10. GRAPHICAL DETERMINATION OF THE DOMINANT PERIODIC SOLUTION FOR THE $1/(s^2-0.4s+1)$ PLANT WITH LINEAR AND PIECEWISE-LINEAR SWITCHING

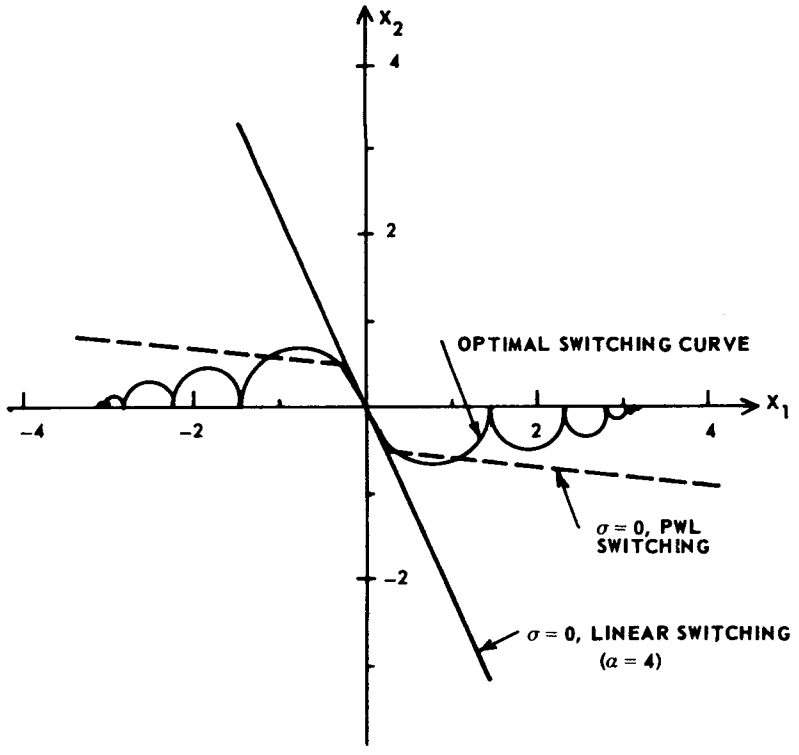


FIG. 11. COMPARISON OF LINEAR AND PIECEWISE-LINEAR SWITCHING LINES WITH THE OPTIMAL SWITCHING LINE, $1/(s^2-0.4s+1)$ PLANT

IV. THIRD-ORDER EXAMPLE

A. STATEMENT OF THE PROBLEM AND OUTLINE OF THE SOLUTION METHOD

As a first example illustrating the design of a PWL switching function and the types of plants and initial conditions for which PWL switching can be expected to give good performance, the plant with transfer function $1/s(s^2+1)$ is considered. Schmidt's design method [Ref. 3] based upon the first switching instant being optimal cannot be applied to this plant because of the undamped oscillatory roots. The design method of Flügge-Lotz and Titus [Ref. 6] was derived specifically for this plant but Kashiwagi [Ref. 20] has shown that the method fails to yield satisfactory responses when the plant poles are moved more than moderately to the left in the s-plane, (say to the left of $\mathrm{Re}\,[s] = -0.3$). Because the plant with all three roots on the imaginary axis (two imaginary and one at the origin) presents a more challenging design problem than the one with well-damped roots, it will be examined first in considerable detail. To verify that the damping of the roots poses no problem to the proposed design methods and to obtain a quantitative comparison of PWL switching functions to the quasi-optimal switching function of Flügge-Lotz and Titus in this situation, the plant with transfer function $1/(s+0.5)(s^2+0.8s+1)$ will be considered briefly at the end of the chapter.

The differential equation of the $1/s(s^2+1)$ plant has the form

$$\dot{\underline{e}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \underline{e} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} u \qquad (4.1)$$

where $\underline{e}^T = (e, \dot{e}, \ddot{e})$ and $|u| = 1$.

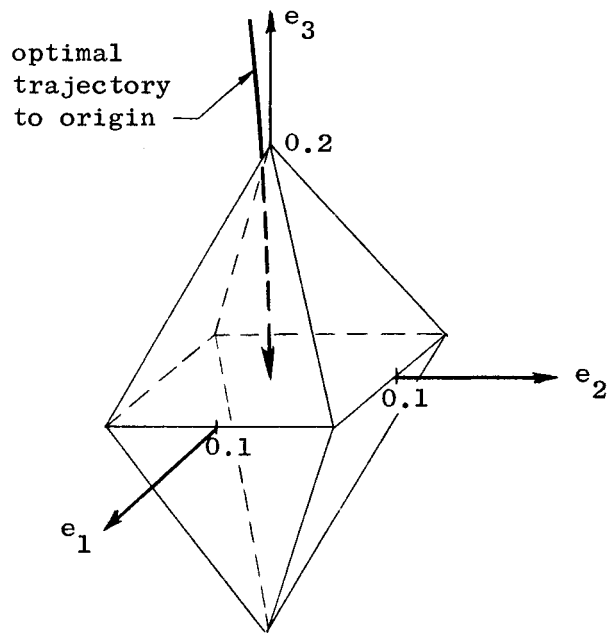
It is assumed that the control system is subject to step-function inputs, which are equivalent to having initial conditions along the e_1 axis. The initial values of e_1 will be taken as the integers $1, 2, \ldots, 8$ and the cost function is given by Eq. (2.2) with $K = 8$, namely,

$$J(\{\rho\}) = \frac{1}{8} \sum_{k=1}^{8} \frac{1}{(T_s)_k} T_k (\{\rho\}, \underline{e}_k^o) \quad . \qquad (4.2)$$

The cost-free region surrounding the origin is defined by the inequalities

$$S = \left\{ \begin{array}{l} |e_1 + 0.5\ e_3| \leq 0.10 \\ |e_1 - 0.5\ e_3| \leq 0.10 \\ |e_2 + 0.5\ e_3| \leq 0.10 \\ |e_2 - 0.5\ e_3| \leq 0.10 \end{array} \right\} .$$

This region, which is sketched below, is the extension to three dimensions of the parallelogram used in the second-order example of Chapter II.



As a consequence of the relatively simple geometry of the trajectories in the canonical space, the optimal times for the initial states along the e_1 axis to reach the origin, $(T_o)_k$, were evaluated by finding a control $\hat{u}(t)$ which transfers the state to the origin and maximizes the Hamiltonian. Since this control is unique for linear plants and the minimum-time criterion,* $\hat{u}(t)$ must be the optimal control. The values of T_o corresponding to the initial conditions used to define J are given below.

e^o	1	2	3	4	5	6	7	8
T_o(sec.)	3.09	3.92	4.56	5.12	5.64	6.14	8.97	9.95

Alternatively, the T_o can be obtained by using the method of Flügge-Lotz and Yin [Ref. 18] . The approximate optimal times to reach the region S, $(T_s)_k$, are found by using Eq. (2.3) with A = 0.20.

The following canonical transformation will be used at times:

$$\underline{x} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underline{e} \quad ; \quad \underline{e} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \underline{x} \; . \qquad (4.3a)\ (4.3b)$$

In the $\underline{x}$ space the differential equation becomes

$$\dot{\underline{x}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \underline{x} + \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} u$$

Examination of the preceding differential equation will show that for constant u the x_1 coordinate changes linearly with time while the x_2 and x_3 coordinates describe circles about the points $x_2 = \pm u$, $x_3 = 0$.

---

*See Chapter 3 of Pontryagin et al. [Ref. 1] .

When linear switching is used, the switching function takes the form of Eq. (3.1) with $n = 3$ and $\rho_3 = 1$, namely,

$$\sigma(\underline{e}) = \rho_1 e_1 + \rho_2 e_2 + e_3 \ .$$

From Eq. (3.5) and a comparison of like powers of s in Eq. (3.4) the two-component linear switching-function parameter vector $\underline{\xi}$ is seen to be

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \omega^2 \\ 2\zeta\omega \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} \ . \qquad (4.4)$$

Having defined both the system and the cost function, the problem is to design a PWL switching function which minimizes the cost $J(\{\rho\})$. The first step is to obtain a preliminary linear switching function by applying the methods described in Chapter III. This design is used to initiate the linear switching performance-surface search on a hybrid computer and it is demonstrated that this surface can be readily searched to obtain the optimal linear switching function. Following this, an optimal PWL switching function will be found by applying the other techniques of Chapter III and searching the resulting performance surface.

B. LINEAR SWITCHING DESIGN GUIDES FOR THE THIRD-ORDER EXAMPLE

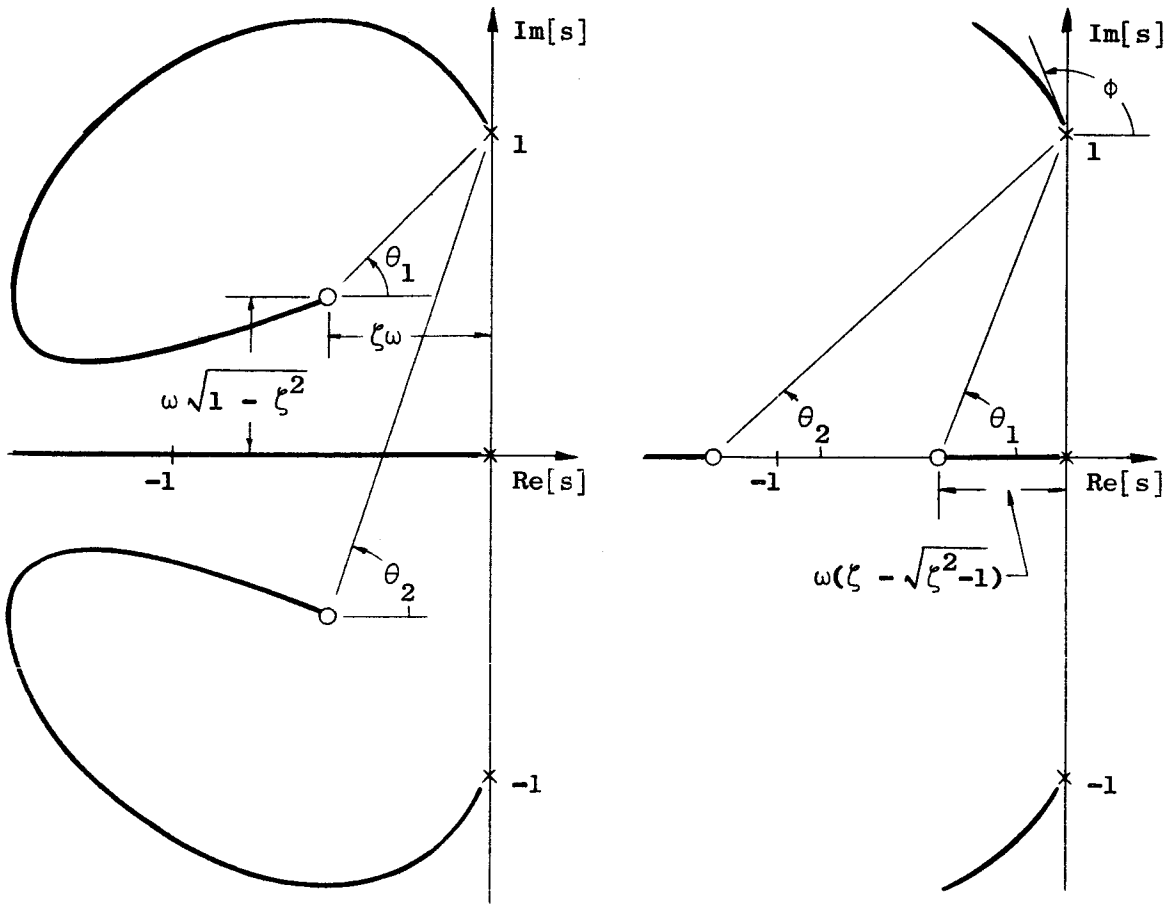
The following three criteria are used in designing the linear switching function:

1. The maximum real part of the switching function zeros is minimized.

2. The angle of departure of the root locus from the complex poles is kept close to 90°.

3. The size of the dominant periodic solution is maximized.

The three design guides used in evaluating the above criteria are evaluated in terms of $\underline{\xi}$ in this section.



a. $0 < \zeta \leq 1$ b. $1 \leq \zeta$

FIG. 12. TYPICAL ROOT-LOCUS PLOTS FOR THE THIRD-ORDER EXAMPLE

1. Maximum Real Part of the Zeros

As shown by the root loci of Fig. 12, the maximum real part of the zeros (denoted by $-\alpha_m$) is given by

$$\alpha_m = \begin{cases} \zeta\omega & 0 < \zeta \leq 1 \\ \omega(\zeta - \sqrt{\zeta^2 - 1}) & 1 \leq \zeta \end{cases} .$$

The line in the $\underline{\xi}$ plane corresponding to $\zeta = 1$ is $\xi_2 = 2\sqrt{\xi_1}$. By using Eq. (4.4) to write the preceding equation in terms of ξ_1 and ξ_2 and then manipulating it so that ξ_2 is expressed as a function of ξ_1 and α_m, the lines in the $\underline{\xi}$ plane of constant α_m are found to be

$$\xi_2 = \begin{cases} 2\alpha_m \\ \frac{1}{\alpha_m}\ \xi_1 + \alpha_m \end{cases} \qquad \xi_1 \geq \alpha_m^2 \ . \tag{4.5}$$

These lines are plotted in Fig. 13a for representative values of α_m.

2. Angle of Departure of the Root Locus from the Complex Poles

The use of the definition of the 180^o root loci and the geometrical relationships shown in Fig. 12 leads to the following expression for the angle of departure ϕ :

$$\phi = \theta_1 + \theta_2 \ . \tag{4.6}$$

If θ_1 and θ_2 are expressed in terms of ζ and ω, the following expression is obtained for ϕ which is valid for all $\zeta > 0$:

$$\phi = \text{arc tan } \frac{2\ \zeta\omega}{\omega^2 - 1} \ .$$

Use of Eq. (4.4) gives the relationship,

$$\phi = \text{arc tan } \frac{\xi_2}{\xi_1 - 1} \ . \tag{4.7}$$

By writing Eq. (4.7) as $\xi_2 = (\xi_1 - 1) \tan \phi$, it is apparent that lines of constant ϕ are merely straight lines passing through the point $\underline{\xi}^T = (1,0)$ with a slope of $\tan \phi$, which implies that the angle between

a line of constant ϕ and the ξ_1 axis is merely ϕ. In Fig. 13b lines of constant ϕ in the $\underline{\xi}$ plane are shown.

3. Size of the Dominant Periodic Solution

It was determined by analog simulation that the dominant periodic solution for the plant under consideration is as shown in Fig. 31, in Appendix A. It is shown that for this periodic solution to exist with linear switching the quarter-period ψ must satisfy Eq. (A6) which becomes

$$\frac{\tan \psi}{\psi} = \frac{\xi_1}{\xi_1 - 1} \quad , \quad \xi_1 > 1 \tag{4.8}$$

when ξ_1 is substituted for ρ_1 and ρ_3 is set equal to unity. As indicated by Fig. 31 the size of the periodic solution is constant for constant ψ which is equivalent to constant ξ_1, provided that $\xi_1 > 1$. Several lines of constant dominant periodic solution size are plotted in Fig. 13c and the corresponding amplitude of the error variable e_3 is given.

It is interesting to note that the two different methods for inferring stability-in-the-large, namely, requiring that $|\phi| > 90^\circ$ and that no periodic solution exist, both yield the same requirement which is that $\xi_1 < 1$.

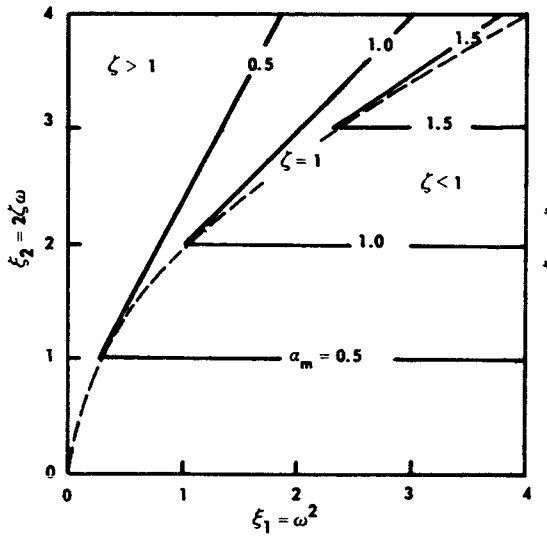
C. DESIGN OF THE OPTIMAL LINEAR SWITCHING FUNCTION FOR THE THIRD-ORDER EXAMPLE

Having accumulated the above qualitative information, it is a relatively simple matter to reduce the likely candidates for the optimal linear switching-function parameters to a relatively small area in the two-dimensional parameter space. From examination of Fig. 13a, it appears reasonable that the parameters should lie somewhere near the line corresponding to $\zeta = 1$, where the fastest rate of decay in chatter is obtained for a given value of ξ_1. Figure 13b indicates that ξ_2 must be raised as ξ_1 is increased in order to keep ϕ from becoming so low that stability becomes a problem, e.g., $\phi < 60^\circ$. Figure 13c

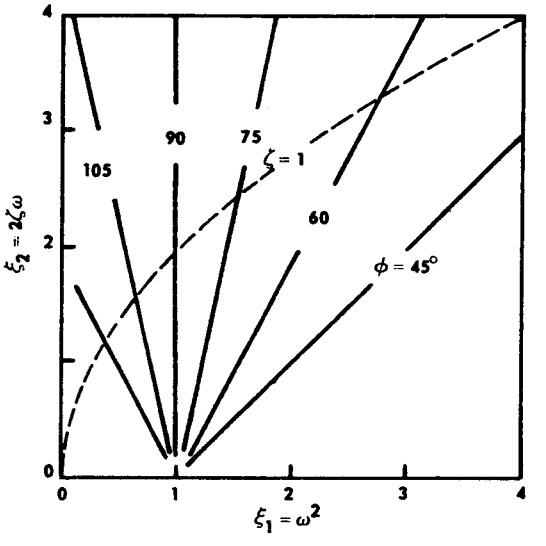
shows that as ξ_1 is increased past $\xi_1 = 1$ the size of the periodic solution, and presumably of the stability region, decreases quite rapidly, independent of ξ_2. This implies that, as ξ_1 and ξ_2 are increased in an effort to increase the rate of decay in chatter and keep ϕ from being too low, there is some value beyond which ξ_1 may not be raised, regardless of any changes in ξ_2.

Should the designer want to choose an initial set of parameters in order to search the performance surface for the optimal linear switching function (for the particular cost function used) a reasonable starting point would be $\underline{\xi}^T = (1,2)$. Here $\phi = 90^o$, $|e_3|_m = \infty$, and α_m takes on its largest value for that particular value of ξ_1. That this is a reasonable choice is shown in Fig. 14 where contours of the actual performance surface defined by Eq. (4.2) are given. It is seen that the suggested initial choice of $\underline{\xi}$, labeled P_1, is quite close to the optimal value of $\underline{\xi}$ and represents a cost of $J = 1.53$ versus $J = 1.20$ for the best linear switching function.

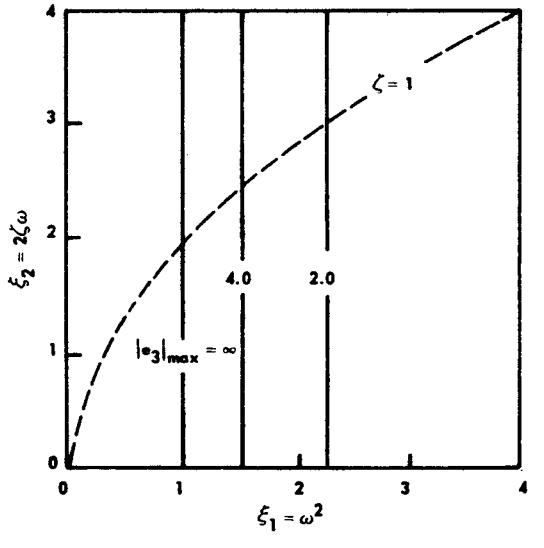
To verify that this choice is feasible, the gradient search process described in Appendix B was used to find the optimal linear switching parameters. However, for the purposes of illustration, the search was initiated from five points (denoted $P_2, P_3, \ldots, P_6$ in Fig. 14) relatively far from the region found by using the qualitative design procedure. From points P_2, P_3, and P_4 the cost was reduced to below 1.25 within six iterations of the search procedure. However, the searches initiated from points P_5 and P_6 both found the relative minima in the region near $\underline{\xi}^T = (0.8, 1.3)$. While the presence of these relative minima may appear to be detrimental to the design procedure, they do not pose a serious problem for three reasons. First, three of the searches found the absolute minimum. Second, the searches were started much further from the absolute minimum than they would have been if the proposed qualitative design information had been used in guiding their selection. Third, on the basis of the second-order results presented in Section II-C, it seems likely that the relative minima would be reduced in size and possibly disappear if more initial conditions were used in the definition of the cost function.



a. Maximum real part of the zeros



b. Angle of departure



c. Dominant Periodic Solution Size

FIG. 13. LINES OF CONSTANT LINEAR SWITCHING DESIGN GUIDES IN THE ξ PLANE, THIRD-ORDER EXAMPLE

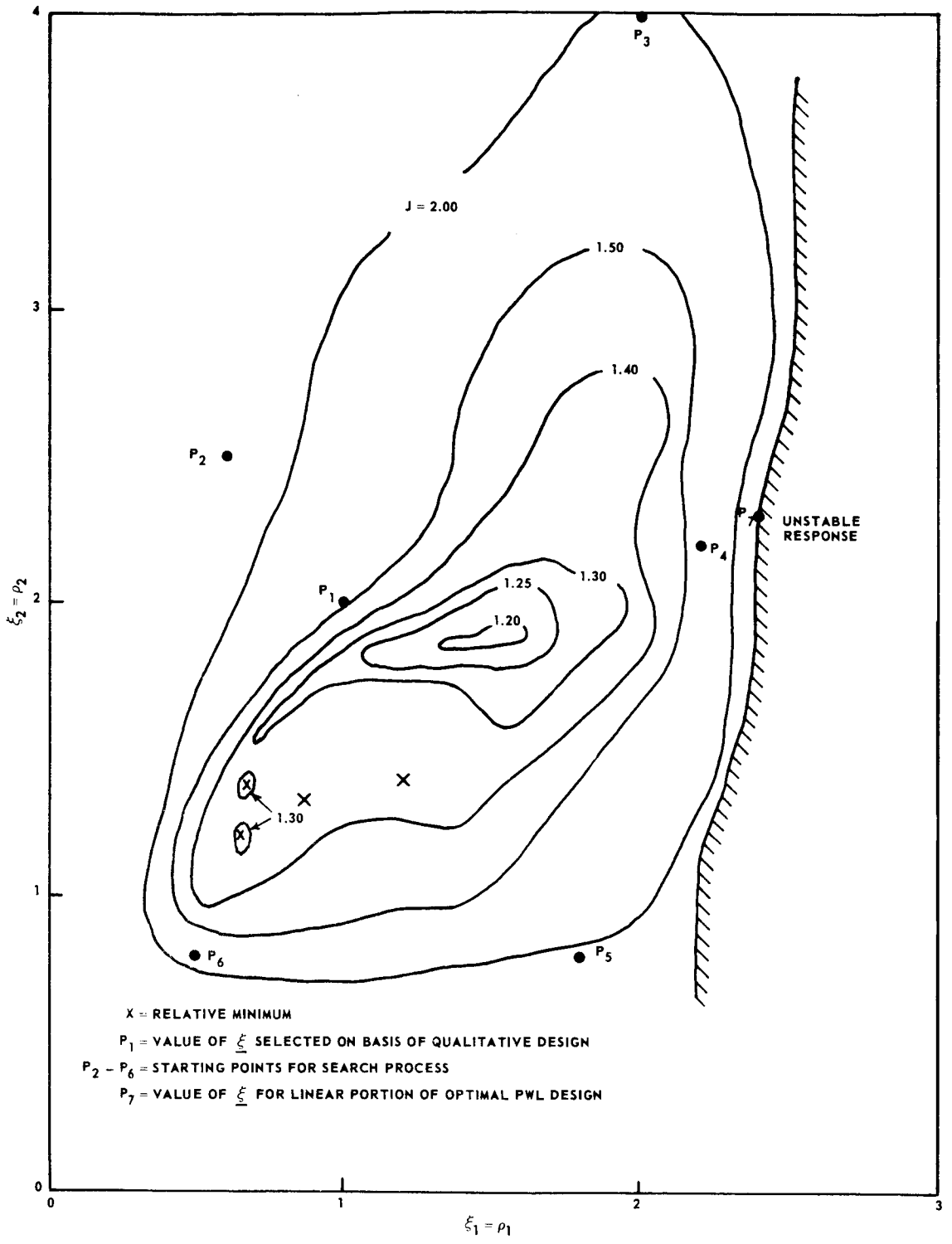


FIG. 14. CONTOURS OF J IN THE $\underline{\xi}$ PLANE FOR LINEAR SWITCHING, THIRD-ORDER EXAMPLE

D. DESIGN OF A PWL SWITCHING FUNCTION FOR THE THIRD-ORDER EXAMPLE

The sub-class of PWL switching functions to be considered is given by Eq. (2.1) and it remains to be decided which of the components $\sigma_i(e_i)$ should be made PWL functions. Because the plant has undamped oscillatory poles it is anticipated that the three-component row vector $\partial\phi/\partial\underline{\rho}$ will provide a basis for this decision, as explained in Section III-C.

By using Eq. (3.6), this vector is readily found to be

$$\frac{\partial\phi}{\partial\underline{\rho}} = \frac{1}{(\xi_1 - 1)^2 + \xi_2^2}\,(-\xi_2,\ \xi_1 - 1,\ \xi_2) \tag{4.9}$$

where the necessary relationships between $\underline{\xi}$, $\underline{\rho}$, and ϕ are given by Eqs. (4.4) and (4.7). Letting $\partial\phi/\partial\rho_i = C_i$, where the C_i are constants, equations for the lines of constant $\partial\phi/\partial\rho_i$ in the $\underline{\xi}$ plane can be obtained.

For example, when $i = 1$, the first component of Eq. (4.9) gives

$$(\xi_1 - 1)^2 + \xi_2^2 + \frac{\xi_2}{C_1} = 0.$$

Completing the square yields

$$(\xi_1 - 1)^2 + \left(\xi_2 + \frac{1}{2C_1}\right)^2 = \left(\frac{1}{2C_1}\right)^2$$

which is the equation of a family of circles with radii $1/2|C_1|$ centered at the points $\underline{\xi}^T = (1, -1/2C_1)$. Similarly, curves of constant C_2 are given by

$$\left[(\xi_1 - 1) - \frac{1}{2C_2}\right]^2 + \xi_2^2 = \left(\frac{1}{2C_2}\right)^2 .$$

Because $\partial\phi/\partial\rho_3 = -\partial\phi/\partial\rho_1$, the curves of constant C_3 are identical to those for $-C_1$. Lines in the $\underline{\xi}$ plane for representative values of $\partial\phi/\partial\rho_i$ are plotted in Fig. 15.

Comparing the lines of constant $\partial\phi/\partial\rho_i$ in Fig. 15 with the contours of constant cost in Fig. 14, it is apparent that in the region of the $\underline{\xi}$ plane where J attains its minimum value $|\partial\phi/\partial\rho_2|$ is considerably smaller than $|\partial\phi/\partial\rho_1|$ and $|\partial\phi/\partial\rho_3|$. Also, since the initial conditions being considered are along the e_1 axis and the e_2 and e_3 coordinates form an oscillatory pair, the e_1 component of the error vector will vary over a considerably greater range than the other two components. One further consideration is the fact that the switching points of the dominant periodic solution depend upon ρ_1 and ρ_3, but not ρ_2, as shown in Appendix A.

For these reasons, the switching function was made PWL by adding one breakpoint to the σ_1 component, yielding the parameter array

$$\{\rho\} = \begin{Bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & x & x \\ 1 & x & x \end{Bmatrix}$$

where the x's denote undefined elements.

The breakpoint was fixed at $\rho_{12} = 1.0$ and the gradient search procedure used to optimize the three slopes ρ_{11}, ρ_{13}, and ρ_{21}, yielding a minimum value of $J = 1.04$ when the gradient search was initiated close to the optimal linear switching parameters where J is evaluated over eight initial conditions, as before. For the reasons given in Section II-D, the random perturbation search procedure was substituted for the gradient method and used to find the minimum cost corresponding to three values of the breakpoint, namely $\rho_{12} = 0.5$, 1.0, and 2.0. For these breakpoints the minimum values of J found were 1.10, 1.04, and 1.17, respectively. With $\rho_{12} = 2.0$ the PWL switching function yielded very little improvement in cost over that obtained with linear switching, presumably because the breakpoint was located so far out

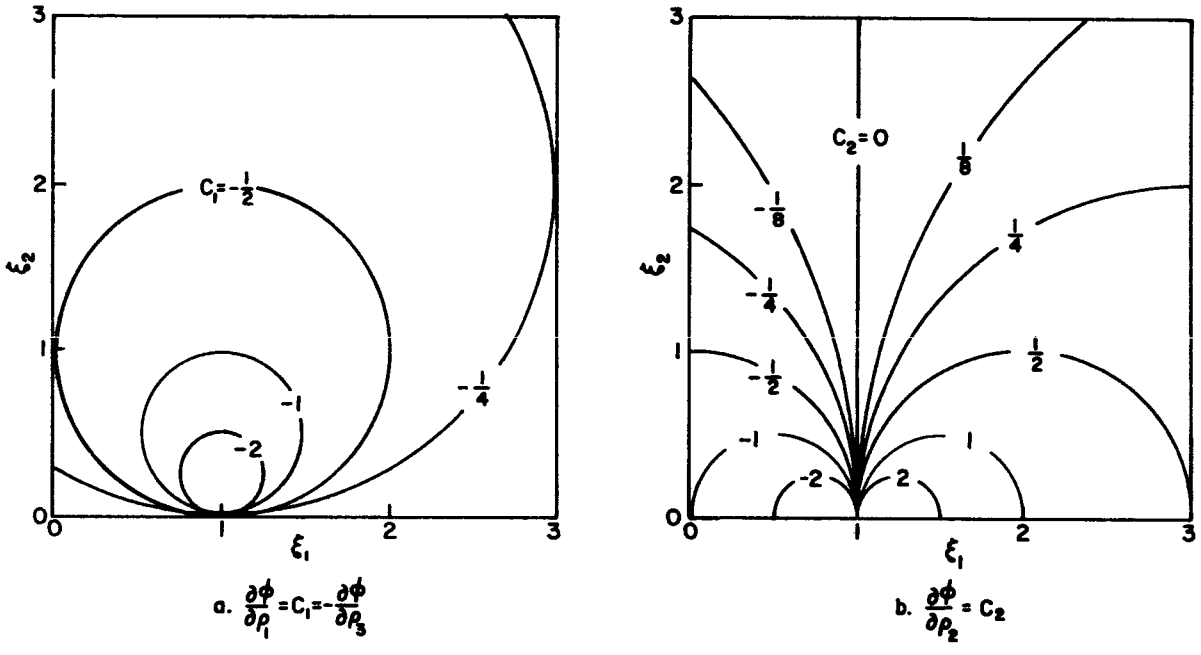


FIG. 15. LINES OF CONSTANT $\partial\phi/\partial\underline{\rho}$ IN THE $\underline{\xi}$ PLANE, THIRD-ORDER EXAMPLE

along the e_1 axis that it was unable to affect the response from more than a portion of those initial conditions used. However, this occurrence would seem to be a property of the particular plant and initial conditions chosen and should not form the basis for any overall conclusions regarding the location of the breakpoints.

By measuring the performance surface in the vicinity of its minimum versus ρ_{11} and ρ_{13} for discrete values of ρ_{21} and with $\rho_{12} = 1.0$, it was found that no relative minima existed in the region measured and that the minimum value of J was indeed 1.04, as was determined from the gradient and random searches.

To investigate the effects of making σ_2 and σ_3 PWL the performance surface was measured for several values of the breakpoint and various values of the three slopes. It was found that making σ_3 PWL, with σ_1 and σ_2 linear, afforded a minimum cost of $J = 1.09$, while making σ_2 PWL gave essentially no improvement over linear switching. The minimum values of cost obtained for these three cases and for linear switching are shown in Table 1, along with their respective optimal parameter values.

TABLE 1. OPTIMAL SWITCHING FUNCTIONS FOR THE $1/s(s^2+1)$ PLANT

No.	Type	I	ρ_{11}	ρ_{12}	ρ_{13}	ρ_{21}	ρ_{22}	ρ_{23}	ρ_{31}	ρ_{32}	ρ_{33}
1	σ_1 PWL	1.04	2.40	1.00	1.10	2.30	x	x	1.00	x	x
2	σ_2 PWL	1.19	1.50	x	x	1.70	1.00	1.90	1.00	x	x
3	σ_3 PWL	1.09	1.95	x	x	2.30	x	x	1.00	0.50	2.45
4	Lin.Sw.	1.20	1.50	x	x	1.90	x	x	1.00	x	x

Note: x denotes an element of $\{\rho\}$ not defined for that particular switching function.

It is interesting to note that the optimal values of ρ_{11} and ρ_{21} when σ_1 is PWL (row 1 in Table 1) correspond to a linear switching function with $\rho_1 = 2.40$ and $\rho_2 = 2.30$. Examination of Fig. 14 shows that this point (denoted P_7) lies on the stability boundary of

the linear switching parameter space, implying that if σ_1 were not PWL at least one of the eight trajectories used in measuring J would be unstable.

In Fig. 16 the system transient responses obtained with those optimal PWL and linear switching functions whose parameters are given by rows 1 and 4 of Table 1 are shown for the eight initial conditions used in defining J. The optimal times for the state to reach the origin are indicated in Fig. 16a and it can be seen that, with the exception of the initial condition $e^o = 6$, the error was reduced to below 0.05 before the optimal time had elapsed. The region S cannot be depicted on Fig. 16; however, $|e|$ must be ≤ 0.1 for the state to be within S. The exact value of e will depend upon the slope and curvature ($\dot{e}$ and $\ddot{e}$) of the transient response. A comparison of the two figures indicates that the only essential difference between the two switching functions is in the small-disturbance response. With the linear switching function the error approaches zero at a somewhat slower rate than it does with the PWL switching function. This is to be expected because, from the root-locus point of view discussed in Chapter III, the zeros corresponding to the linear switching function must be located further to the right in the s-plane than those corresponding to the linear portion of the PWL switching function in order to insure stability for large states. For all initial conditions in the range $1 \leq e^o \leq 8$, the maximum overshoot was 0.02 for the PWL switching function and 0.04 for the linear switching function.

E. STABILITY CONSIDERATIONS FOR THE THIRD-ORDER EXAMPLE

Having found that by making σ_1 PWL a minimum cost of $J = 1.04$ can be obtained, one might ask what is to be gained by making one or more of the other components of σ be PWL. Since $J = 1.00$ when each of the K trajectories is optimal, it is apparent that the cost of $J = 1.04$ cannot be reduced much further, regardless of what switching function is used. However, when σ_1 is the only PWL function and $\{\rho\}$ is adjusted to attain this value of cost, it can be shown that $\{\hat{\rho}\}$, the optimal value of $\{\rho\}$, is extremely close to the region in the

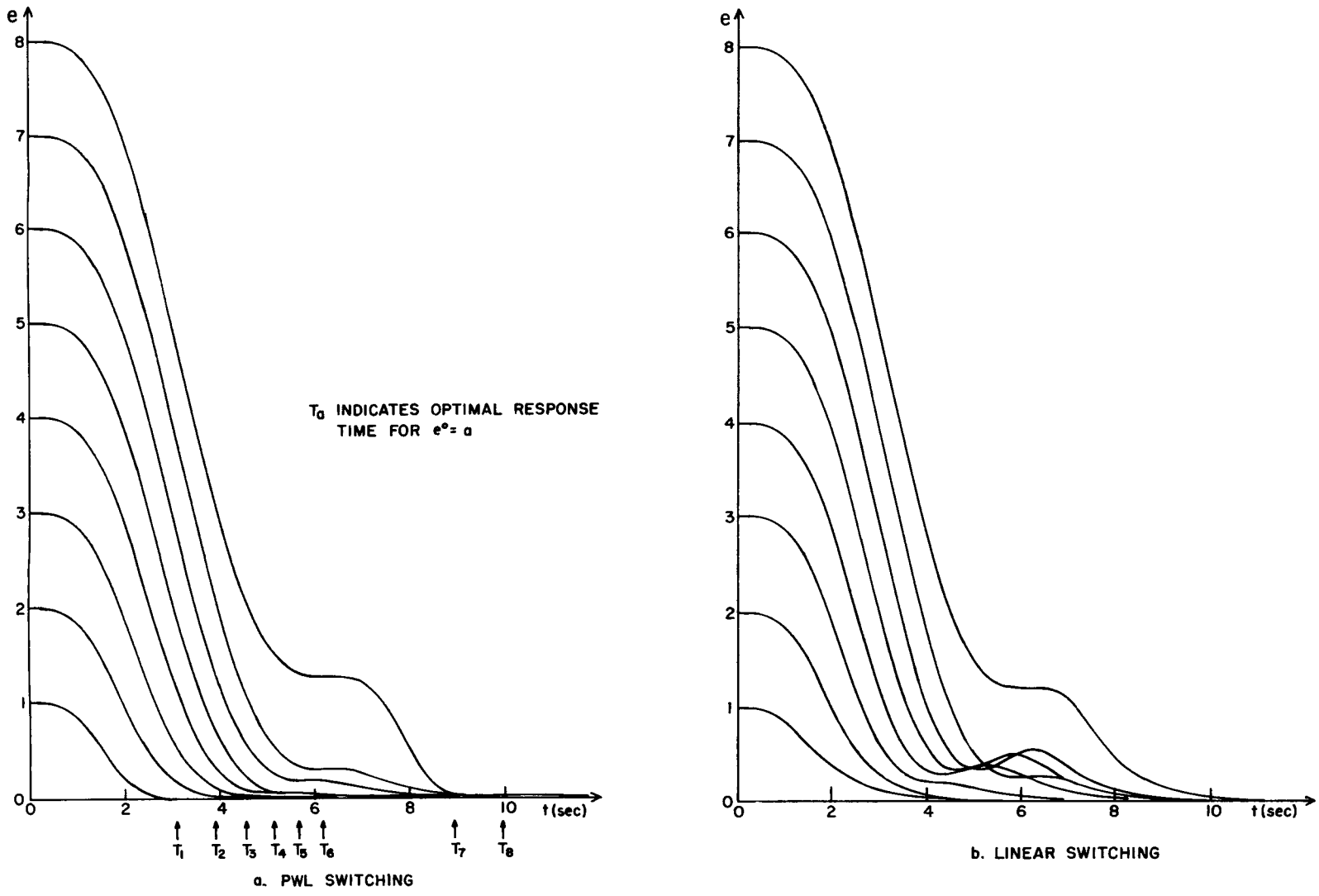


FIG. 16. STEP-FUNCTION RESPONSES WITH OPTIMAL PIECEWISE-LINEAR AND LINEAR SWITCHING FUNCTIONS, THIRD-ORDER EXAMPLE

parameter space corresponding to one or more unstable trajectories among those used to define J.

Figure 17 shows the stability boundary in the ρ_{11}, ρ_{13} plane for $\rho_{12} = 1.0$ and $\rho_{21} = 2.30$, as obtained by simulating the system on the analog computer in the repetitive mode. The region to the left of the stability boundary corresponds to PWL switching functions for which all initial conditions along the e_1 axis between $0 \leq e_1^o \leq 8$ are stable and the region to the right corresponds to switching functions yielding an unstable trajectory from at least one initial condition in this range. The point in the parameter space corresponding to $\{\hat{\rho}\}$ is indicated in Fig. 17 and it can be seen that it lies relatively close to the stability boundary. The situation shown is undesirable in a practical control system because relatively small fluctuations in the controller or plant parameters could result in unstable trajectories.

In addition to the stability boundary in the parameter space it is of substantial engineering interest to consider the stability boundary in the state-space. When the PWL switching function is described by $\{\hat{\rho}\}$, the stability boundary in the first quadrant of the e_1, e_2 plane is as shown in Fig. 18. In other words, any initial condition for which $0 \leq e_1^o \leq 8$, $e_2^o > 1$, and $e_3^o = 0$ will be unstable. This situation also poses severe drawbacks from an engineering point of view because of the disastrous consequences of initial conditions for which $\dot{e}$ and $\ddot{e}$ are not very small.

It was mentioned in Section III-B that during numerous analog simulations the existence of the dominant periodic solution shown in Fig. 31 has been observed to play a vital role in the stability of the system. As shown in Appendix A, a condition necessary for the existence of this periodic solution is that the switching surface, i.e., all points where $\sigma = 0$, intersect the locus of possible switching points shown in Fig. 33. Since this locus lies entirely in the e_1, e_3 plane, only the intersection of the switching surface with this plane need be considered. If σ_1 and σ_3 are both PWL with one breakpoint apiece, as shown in Fig. 19, the intersection of the switching surface with the e_1, e_3 plane will be as shown in Fig. 20.

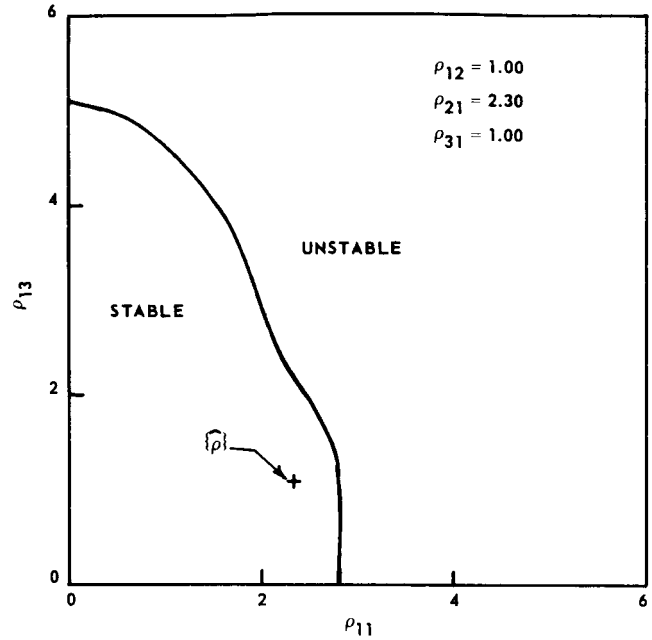


FIG. 17. STABILITY REGION IN THE ρ_{11}, ρ_{13} PLANE FOR σ_1 PIECEWISE-LINEAR AND INITIAL STATES ALONG THE e_1 AXIS, THIRD-ORDER EXAMPLE

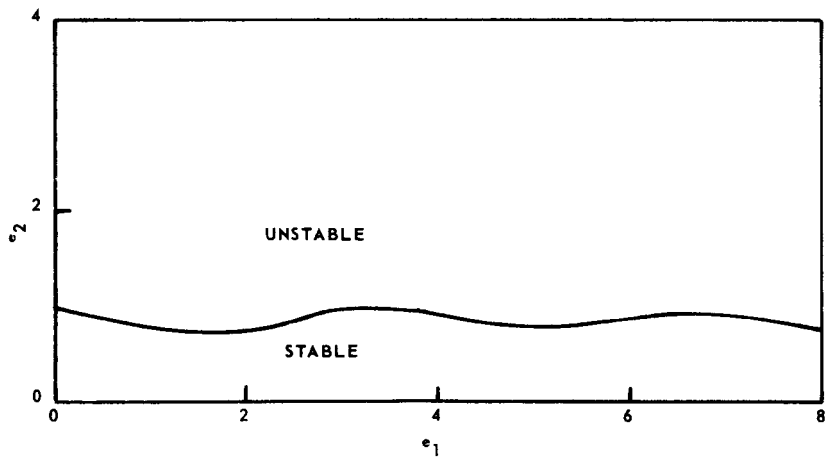


FIG. 18. STABILITY REGION IN THE e_1, e_2 PLANE FOR σ_1 PIECEWISE-LINEAR, THIRD-ORDER EXAMPLE

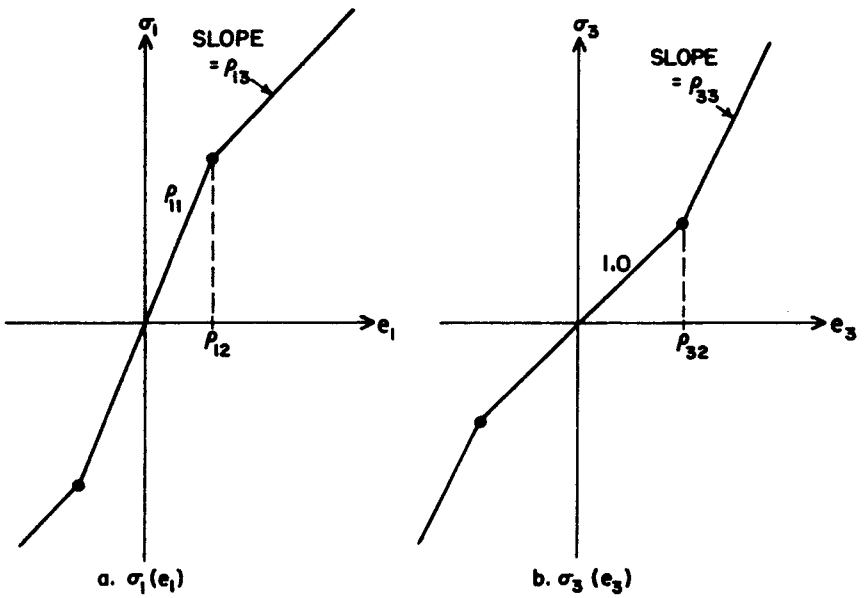


FIG. 19. $\sigma_1(e_1)$ AND $\sigma_3(e_3)$ FOR A TYPICAL PIECEWISE-LINEAR SWITCHING FUNCTION, THIRD-ORDER EXAMPLE

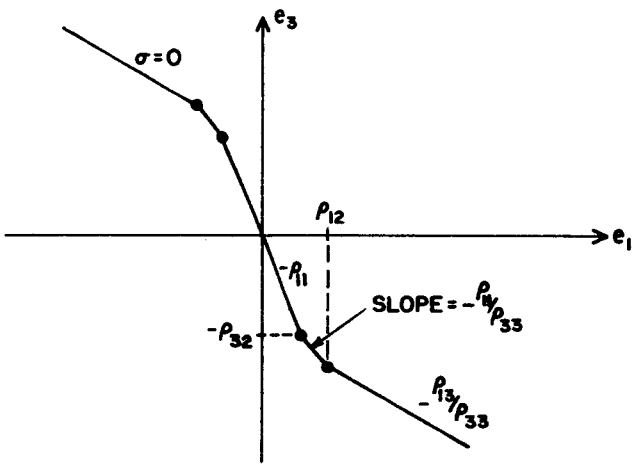


FIG. 20. INTERSECTION OF THE SWITCHING SURFACE WITH THE e_1, e_3 PLANE FOR $\sigma_1(e_1)$ AND $\sigma_3(e_3)$ OF FIG. 19

Now it is an easy matter to evaluate the effect upon the dominant periodic solution of making σ_3 be PWL when σ_1 is fixed so as to yield the minimum cost, i.e., $\rho_{11} = 2.40$, $\rho_{12} = 1.00$, and $\rho_{13} = 1.10$ (see row 1 of Table 1). When σ_3 is linear, with $\rho_{31} = 1.00$, the intersection of the switching surface with the e_1, e_3 plane is the line marked $\rho_{32} = \infty$ in Fig. 21. It is apparent from the intersection of this line and the locus of possible periodic-solution switching points at the points N_1 and N_2 in Fig. 21 that the condition on the switching points necessary for the existence of the periodic solution has been satisfied.* Therefore, when only σ_1 is PWL with $\rho_{12} = 1.00$, the periodic solution exists exactly as if the switching function were linear. This situation accounts for the small margin of stability both in the parameter space for initial conditions along the e_1 axis and in the state-space for initial conditions not along the e_1 axis.

If σ_3 is made PWL with $\rho_{32} = 2.00$ and $\rho_{33} = 2.00$, the intersection of the switching surface with the e_1,e_3 plane becomes the line in Fig. 21 denoted by $\rho_{32} = 2.00$. It can be seen that the switching surface intersection line and the locus of possible switching points become tangent to one another but, with the trivial exception of the origin, do not intersect. That $\rho_{32} = 2.00$ marks the transition between two distinct types of behavior can be seen by examination of the line marked $\rho_{32} = 1.00$. Here, the switching surface has no non-trivial intersection with the locus of possible dominant periodic solution switching points. Furthermore, the maximum value attained by $|e_3|$ in any of the trajectories defining the cost function is 1.00, so the addition of the breakpoint and increase of slope at $\rho_{32} = 1.00$ will have no effect upon the minimum cost obtainable for the initial conditions comprising J. Also, as long as $\rho_{13}/\rho_{33} < 1$ the slope of the switching surface intersection in the e_1, e_3 plane will be greater than -1 (see Fig. 20) and there can be no intersection of the

---

*The sufficiency of this condition has been verified by analog simulation.

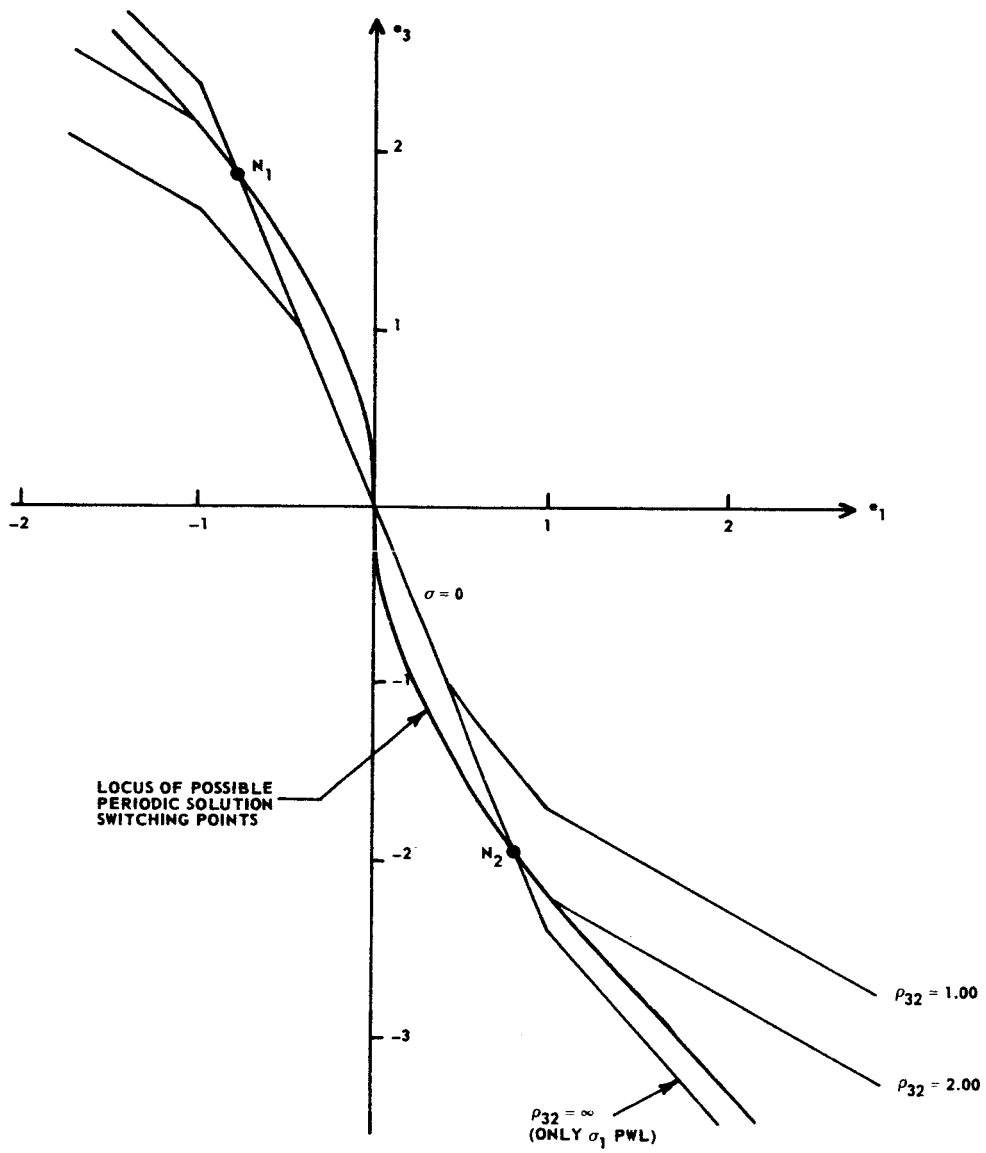


FIG. 21. LOCUS OF POSSIBLE PERIODIC SOLUTION SWITCHING POINTS AND INTERSECTIONS OF PIECEWISE-LINEAR SWITCHING SURFACES WITH THE e_1, e_3 PLANE, THIRD-ORDER EXAMPLE

switching surface and the locus of possible periodic solution switching points for values of $|e_1|$ larger than those shown in Fig. 21.

The effect upon the stability region in the parameter space of making σ_3 PWL with $\rho_{33} = 2.00$ and with various values of the break-point ρ_{32} can be seen in Fig. 22. The minimum cost was $I = 1.04$ for $\rho_{32} \geq 1.00$. Also, it was observed that the point of minimum cost in the ρ_{11}, ρ_{13} plane, denoted $\{\hat{\rho}\}$, remained essentially unchanged when σ_3 was made PWL. The effect upon the stability region in the state-space of making σ_3 PWL with $\rho_{32} = 1.00$ and $\rho_{33} = 2.00$ was evaluated on the analog computer. It was found that the system was asymptotically stable for all initial conditions in the region $\|\underline{e}\| \leq 100$. It was not possible to attain larger initial states due to scaling limitations, but it was observed that the character of the response was essentially independent of the size or location of the initial state, as long as it was large, i.e., $\|\underline{e}\| > 10$.

Therefore, if the switching function described by the array

$$\{\rho\} = \begin{Bmatrix} 2.40 & 1.00 & 1.10 \\ 2.30 & x & x \\ 1.00 & 1.00 & 2.00 \end{Bmatrix} \tag{4.10}$$

is used, the resulting system, which is shown in block-diagram form in Fig. 23, will have a cost of four percent above optimal for the cost function defined by Eq. (4.2), step-function responses as shown in Fig. 16a, and will be asymptotically stable at least in the region $\|\underline{e}\| \leq 100$.

From a performance surface searching point of view, the region of stability in the state-space could also be enlarged by using a wider variety of initial conditions to define the performance surface which, in turn, will result in a new value of $\{\hat{\rho}\}$. If the initial conditions chosen were not typical of the operation of the system with respect to e_2 and e_3, then others could have been added. The choice of the switching function components to be made PWL and of the initial values for the PWL parameters to be searched would be carried out as in the previous example.

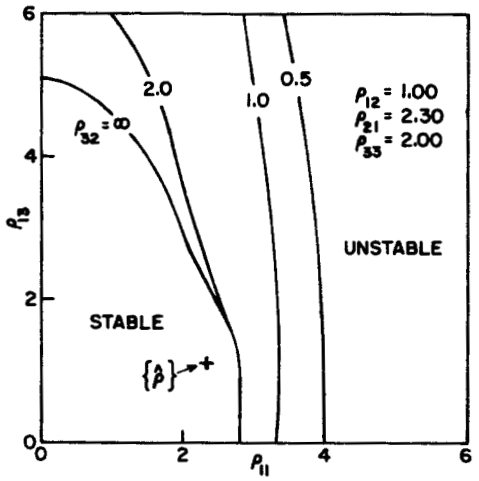


FIG. 22. STABILITY REGIONS IN THE ρ_{11}, ρ_{13} PLANE FOR VARIOUS VALUES OF ρ_{32} WITH BOTH σ_1 AND σ_3 PIECEWISE-LINEAR AND INITIAL STATES ALONG THE e_1 AXIS, THIRD-ORDER EXAMPLE

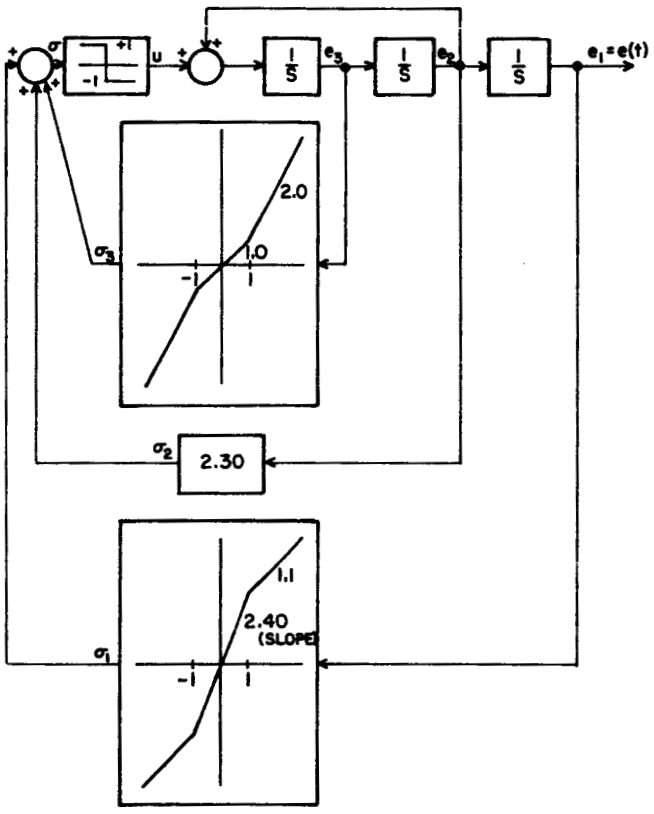


FIG. 23. BLOCK-DIAGRAM OF THE $1/s(s^2+1)$ PLANT WITH THE PIECEWISE-LINEAR SWITCHING FUNCTION CORRESPONDING TO THE ARRAY (4.10)

F. QUALITATIVE DESIGN OF LINEAR AND PWL SWITCHING FUNCTIONS FOR A THIRD-ORDER PLANT WITH DAMPED ROOTS

To show that plants having well-damped roots present no difficulties for the qualitative design procedures given, sub-optimal linear and PWL switching functions are found for the $1/(s + 0.5)(s^2 + 0.8\ s + 1)$ plant. It is concluded that, when the plant roots are well-damped and the cost-free region is of moderate size (say $\|\underline{x}\| \leq 0.20$ for this plant,* if $|u| = 1$) linear switching can provide close-to-optimal response and making the switching function PWL cannot provide any significant improvements in the cost or the region of stability. Likewise, the response times are considerably less sensitive to the switching function parameters than when the plant poles are undamped. Only if the cost-free region is quite small (say $\|\underline{x}\| \leq 0.02$ for this plant) is the use of PWL switching warranted.

By applying the qualitative design criteria stated in Section B four trial linear switching functions were selected and the trajectories from four widely spaced initial conditions were simulated on the analog computer. The response times obtained with the switching function yielding the lowest total response time, i.e.,

$$\underline{\rho}^T = (1,2,1) \ , \tag{4.11}$$

are shown in Table 2 with the corresponding initial conditions and optimal times. The transient responses are given in Fig. 24a. Kashiwagi [Ref. 20] has found that the quasi-optimal switching function of Flügge-Lotz and Titus [Ref. 6] reduced initial condition No. 1 to

---

*The canonical variable $\underline{x}$ is as defined by Flügge-Lotz and Titus [Ref. 6]. For this plant, one obtains

$$\underline{x} = \begin{pmatrix} 1.000 & 0.800 & 1.000 \\ 0.500 & 1.200 & 0.400 \\ 0 & 0.458 & 0.916 \end{pmatrix} \underline{e} \ .$$

$\|\underline{x}\| \leq 0.20$ in 12.01 sec., versus 8.20 sec. for the linear switching function given above.

TABLE 2. LARGE-DISTURBANCE RESPONSE TIMES FOR THE $1/(s+0.5)(s^2+0.8s+1)$ PLANT WITH LINEAR SWITCHING GIVEN BY EQ. (4.11)

Initial Condition				Response Time (sec.)	
No.	e^o	$\dot{e}^o$	$\ddot{e}^o$	Optimal	Linear Switching
1	-27.7	5.5	23.2	7.10	8.20
2	15.6	9.6	-4.0	6.12	7.05
3	0.9	1.6	-9.1	5.82	6.35
4	-14.9	6.4	7.5	5.96	7.10

Note: The optimal response time is to the origin; the sub-optimal times are to $\|\underline{x}\| \leq 0.20$.

Efforts to improve upon the results given in Table 2 by making the switching function PWL and using the qualitative criteria given in Section III-C to select several sets of parameter values proved unsuccessful because the breakpoints of the PWL σ_i were close to or within the cost-free region, implying that the PWL switching function was essentially linear. To show the utility of PWL switching when the cost-free region is small, a switching function was designed using a set of four initial conditions having $\|\underline{x}^o\| \simeq 2.0$ and a cost-free region defined by $\|\underline{x}\| \leq 0.02$. This switching function, described by the parameter array

$$\{\rho\} = \begin{Bmatrix} 6.00 & 0.05 & 2.00 \\ 3.00 & x & x \\ 1.00 & x & x \end{Bmatrix}, \qquad (4.12)$$

yields the response times given in Table 3 and the transient responses shown in Fig. 25b for the four small initial conditions.

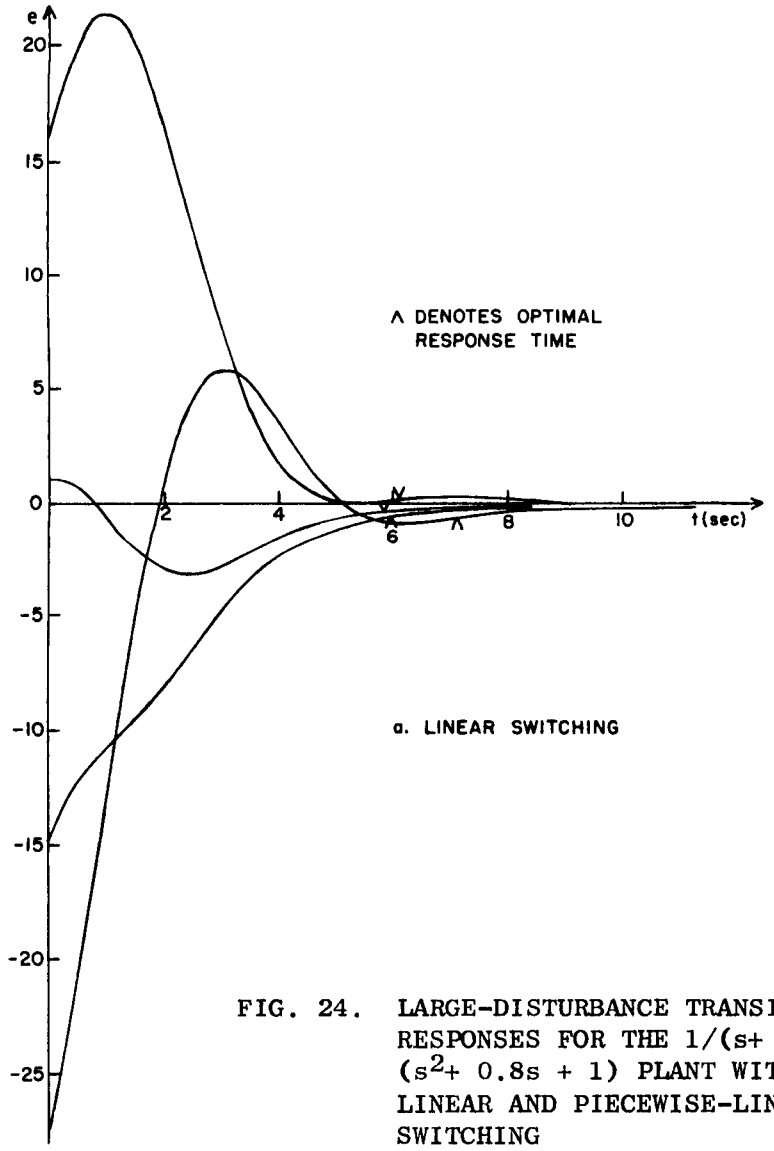
TABLE 3. SMALL-DISTURBANCE RESPONSE TIMES FOR THE $1/(s+0.5)(s^2+0.8s+1)$ PLANT WITH PWL SWITCHING GIVEN BY EQ. (4.12) AND LINEAR SWITCHING GIVEN BY EQ. (4.11)

Initial Condition				Response Time (sec.)		
No.	e^o	$\dot{e}^o$	$\ddot{e}^o$	Optimal	PWL Sw.	Lin. Sw.
5	0.11	-0.57	1.56	1.40	2.0	2.3
6	-0.26	-0.39	1.81	1.95	3.7	4.0
7	-0.99	0.74	0.79	1.99	3.7	4.8
8	-1.16	1.66	-0.85	1.77	3.6	5.0

Note: The optimal response time is to the origin; the sub-optimal times are to $\|\underline{x}\| \leq 0.02$.

In order to substantiate further the conclusions stated at the beginning of this section, the PWL switching function (4.12) was used to control the large initial conditions (1,2,3, and 4), yielding the transient responses shown in Fig. 24b. Because the breakpoint is so small ($\rho_{12} = 0.05$) the PWL switching function is essentially a linear switching function with $\underline{\rho}^T = (2, 3, 1)$ for all but small states. Since these effective linear switching coefficients are similar to those given by Eq. (4.11) it is not surprising that the transient response of Figs. 24a and b are similar. Finally, the small-disturbance responses of the linear switching function are given in Fig. 25a and the response times in Table 3.

It is readily apparent that the design based on large initial conditions and a relatively large cost-free region has resulted in small transients which decay very slowly. If the performance corresponding to either switching function is not sufficiently close to the optimal, the parameter values given can be used to initiate an optimization procedure.



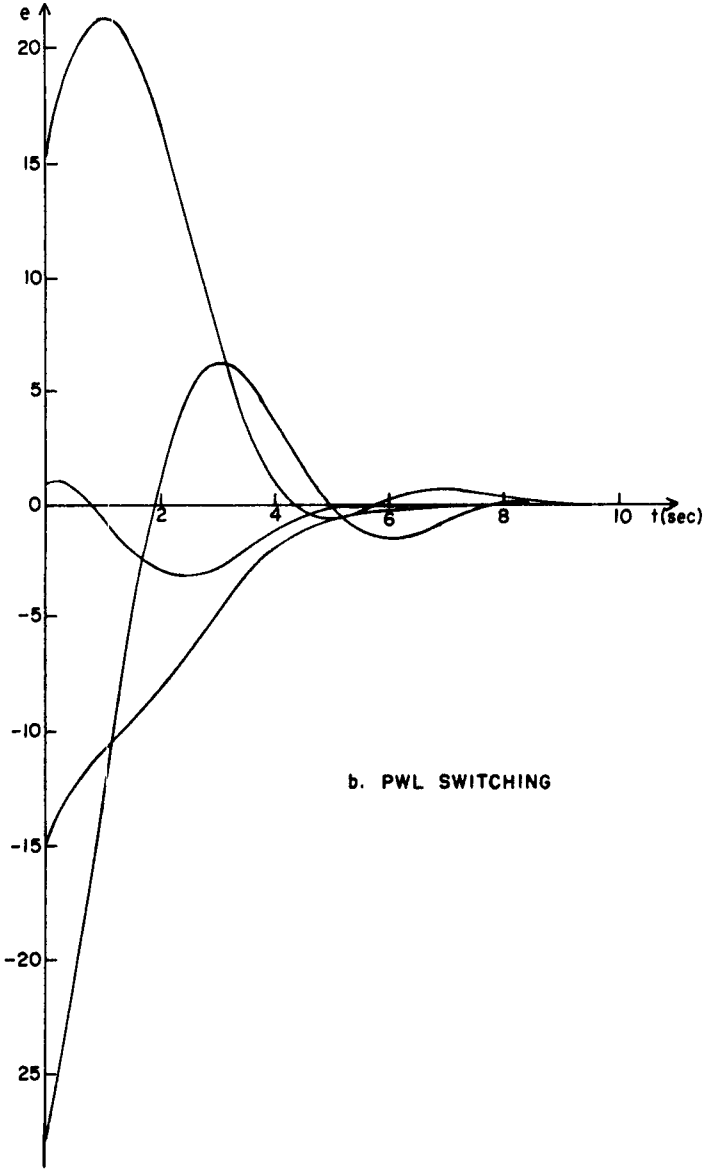
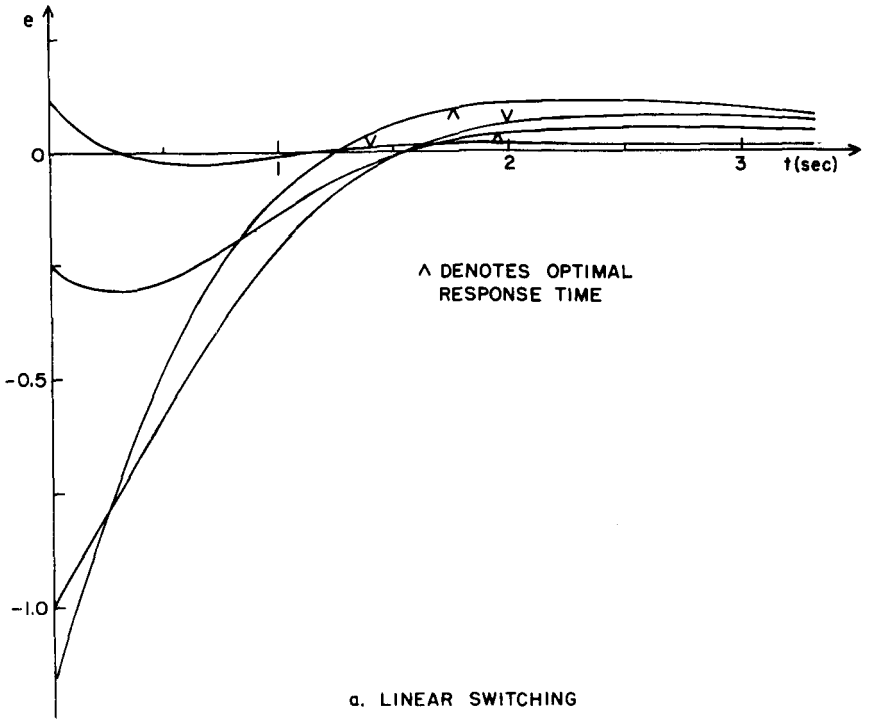


FIG. 24. LARGE-DISTURBANCE TRANSIENT RESPONSES FOR THE $1/(s+0.5)(s^2+0.8s+1)$ PLANT WITH LINEAR AND PIECEWISE-LINEAR SWITCHING



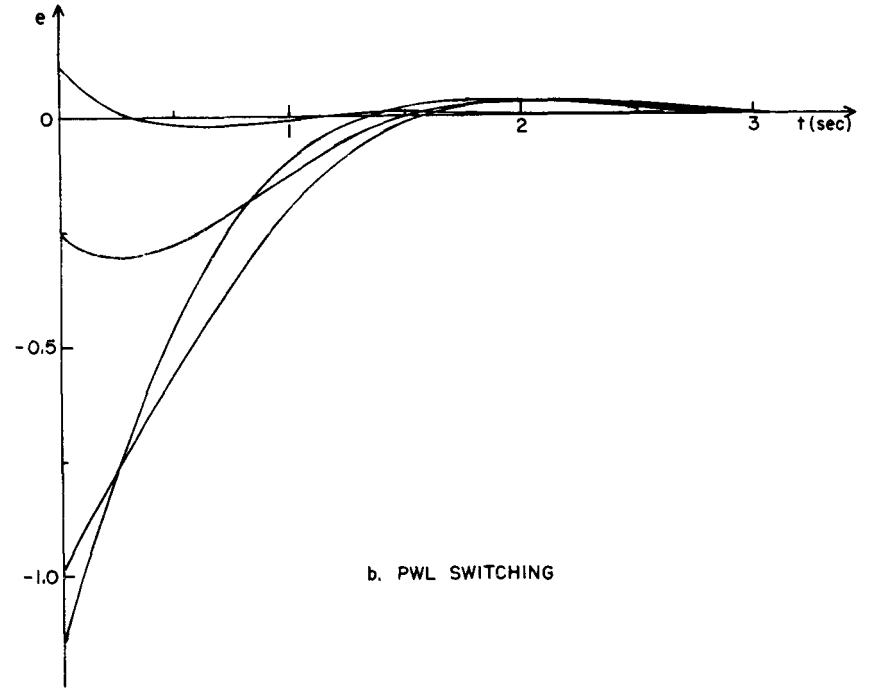


FIG. 25. SMALL-DISTURBANCE TRANSIENT RESPONSES FOR THE $1/(s + 0.5)(s^2 + 0.8s + 1)$ PLANT WITH LINEAR AND PIECEWISE-LINEAR SWITCHING

V. FOURTH-ORDER EXAMPLE

A. PROBLEM STATEMENT

The switching surfaces designed for a third-order control system in the previous chapter illustrated the use of the design guides and surface searching procedures and raised many of the problems to be encountered in their application. Since one of the main reasons for applying the root-locus method and considering the dominant periodic solution is to reduce the complications involved in control design for higher-order systems, it is of interest to show that these design methods are applicable to a fourth-order (and presumably higher-order) plant with relatively little additional complexity.

Also, it will indeed become apparent that the design of an optimal PWL switching function by surface searching methods becomes more complicated as the order of the system increases, due mainly to the increased dimensionality of the performance surface to be searched. It will be shown, however, that the qualitative design methods are extremely helpful in finding a simple sub-optimal PWL switching function which yields good performance over a wide range of initial conditions which could serve either as a final design or as the initial point of a parameter optimization.

The plant with transfer function $1/s^2(s^2+1)$ will be considered in the same manner as in the previous chapter and it will be found that much of the design process will be an extension of work already done for the third-order example. As mentioned in the previous chapter, the results extend immediately to fourth-order systems with roots in the LHP.

The plant and contactor are described by the differential equation

$$\dot{\underline{e}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \underline{e} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} u \qquad (5.1)$$

where $\underline{e}^T = (e, \dot{e}, \ddot{e}, \dddot{e})$ and $|u| = 1$.

The cost function given by Eq. (4.2) will be used, with the initial states now being of the form $(\underline{e}^o)^T = (e_1^o, 0, 0, 0)$, where $e_1^o = 1, 2, \ldots, 8$. The optimal response times corresponding to the initial conditions used to define J are obtained as in the third-order example of Chapter IV, i.e., by using the simple geometry of the trajectories in the canonical space and the uniqueness of the optimal control, and are given below.

e^o	1	2	3	4	5	6	7	8	16	32
T_o (sec.)	4.28	5.05	5.56	5.97	6.30	6.60	6.87	7.12	8.73	11.38

The cost-free region S is defined by the inequalities

$$S = \left\{ \begin{array}{c} |e_1 + e_2| \leq 0.10 \\ |e_1 - e_2| \leq 0.10 \\ |e_3 + 0.5\, e_4| \leq 0.10 \\ |e_3 - 0.5\, e_4| \leq 0.10 \end{array} \right\}$$

To simplify the visualization of the trajectories, the canonical transformation

$$\underline{x} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underline{e} \quad ; \quad \underline{e} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \underline{x}, \qquad (5.2a) \quad (5.2b)$$

which will transform Eq. (5.1) into

$$\dot{\underline{x}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \underline{x} + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} u,$$

is defined. Since the state-space is four-dimensional, the trajectories will be represented by their projections on the x_1, x_2 plane which are parabolas and on the x_3, x_4 plane which are arcs of circles centered at the points $x_3 = \pm u$, $x_4 = 0$.

The linear switching function, expressed in the error variables, is

$$\sigma(\underline{e}) = \rho_1 e_1 + \rho_2 e_2 + \rho_3 e_3 + e_4$$

where ρ_4 has been set equal to unity. The three-dimensional linear switching-function parameter space is described by the vector $\underline{\xi}$ whose components are given by Eq. (3.5) with $n = 4$, i.e.,

$$\underline{\xi}^T = (\omega^2, 2\zeta\omega, \alpha). \tag{5.3}$$

By equating like powers of s in Eq. (3.4), the linear switching coefficient vector $\underline{\rho}$ can be expressed in terms of the components of $\underline{\xi}$ as

$$\underline{\rho}^T = (\xi_1\xi_3, \xi_1 + \xi_2\xi_3, \xi_2 + \xi_3, 1). \tag{5.4}$$

While $\underline{\rho}$ is easily found knowing $\underline{\xi}$, the converse requires the solution of two cubic equations whose roots are not even unique when $\zeta > 1$. For this reason, and the fact that most of the relationships to be found are conveniently expressed in terms of its components, $\underline{\xi}$ will be used throughout most of the analysis, rather than $\underline{\rho}$. When $n \leq 3$ and the state variables are e and its derivatives, $\underline{\xi}$ and the first $(n-1)$ components of $\underline{\rho}$ are identical.

A single relationship is established between the components of $\underline{\xi}$ when one of the qualitative design guides is held constant, thereby defining a surface in the three-dimensional $\underline{\xi}$ space upon which that particular design guide is constant. For this example lines of constant value for the qualitative design guides will be plotted versus ξ_1 and ξ_2 for several fixed values of ξ_3. For a fifth-order plant several values could be assumed for the damping ratios of the two pairs of switching function zeros, yielding a two-dimensional plane in which the qualitative design guides could be plotted.

The requirement that the switching function zeros lie in the LHP can be found by applying the Routh-Hurwitz criterion and is that $\rho_2\rho_3 > 1$, in addition to the obvious requirement that the $\rho_j > 0$, $j = 1,2,3$. In terms of the $\underline{\xi}$ space, the requirement is simply that the $\xi_i > 0$, $i = 1,2,3$.

B. LINEAR SWITCHING DESIGN GUIDES FOR THE FOURTH-ORDER EXAMPLE

The design guides to be used in applying the three linear switching design criteria stated in Section IV-B are evaluated in terms of $\underline{\xi}$.

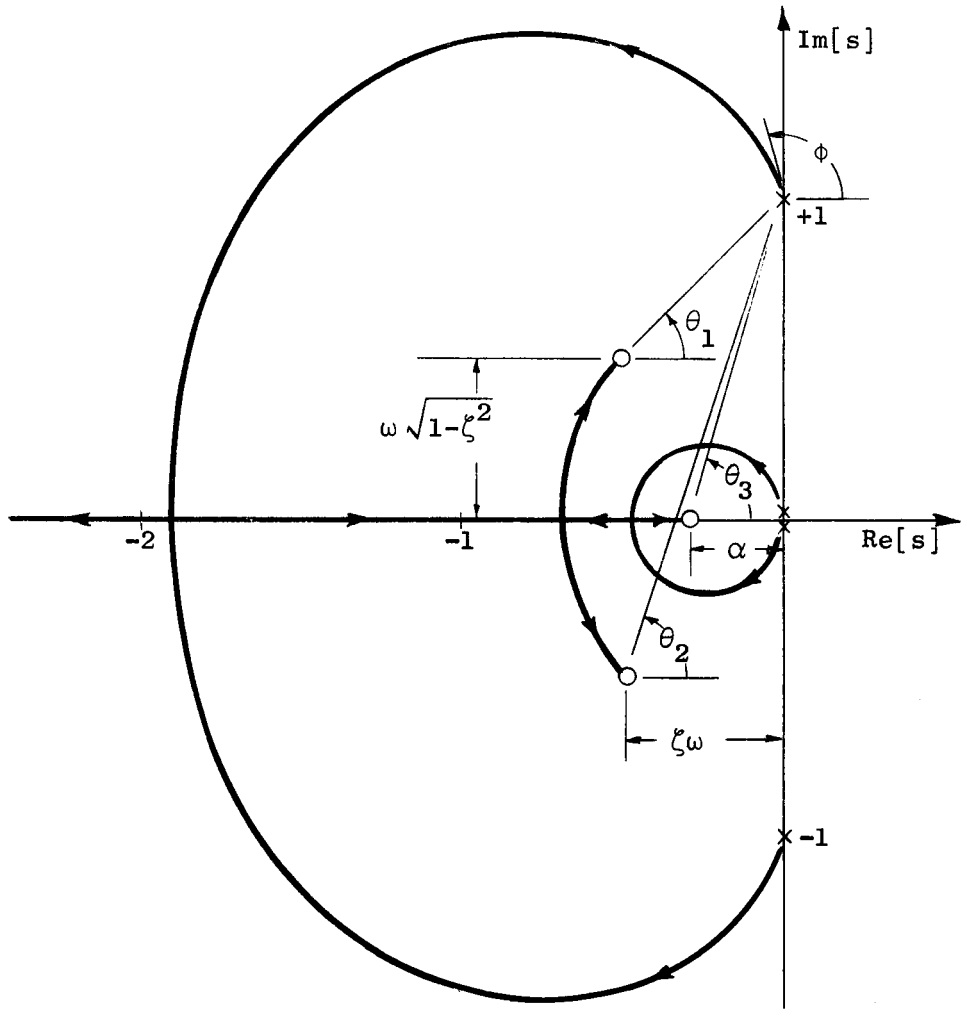


FIG. 26. TYPICAL ROOT-LOCUS PLOT FOR THE FOURTH-ORDER EXAMPLE

1. Maximum Real Part of the Zeros

When $0 < \zeta \leq 1$, examination of the root-locus plot in Fig. 26 reveals that the maximum real part of the zeros (denoted by $-\alpha_m$) is given by

$$\alpha_m = \min [\alpha, \zeta\omega] \qquad 0 < \zeta \leq 1.$$

If $\zeta \geq 1$, the maximum real part of the two zeros described by ζ and ω is as shown in Fig. 12b and

$$\alpha_m = \min [\alpha, \omega(\zeta - \sqrt{\zeta^2-1})] \qquad 1 \leq \zeta.$$

By using Eq. (5.3) to express the two equations above in terms of $\underline{\xi}$, the following set of relationships defining a surface of constant α_m in the $\underline{\xi}$ space can be derived in a manner analogous to that of Section IV-B-1:

$$\begin{aligned}
&\text{(i)} \quad \xi_1 \geq \alpha_m^2, \quad 2\alpha_m \leq \xi_2 \leq \xi_1/\alpha_m + \alpha_m, \quad \xi_3 = \alpha_m \\
&\text{(ii)} \quad \xi_1 \geq \alpha_m^2, \quad \xi_2 = 2\alpha_m, \quad \xi_3 \geq \alpha_m \\
&\text{(iii)} \quad \xi_1 \geq \alpha_m^2, \quad \xi_2 = \xi_1/\alpha_m + \alpha_m, \quad \xi_3 \geq \alpha_m .
\end{aligned} \tag{5.5}$$

Each of the above relationships describes a portion of a plane and the three planes intersect at the point $\underline{\xi}^T = (\alpha_m^2, 2\alpha_m, \alpha_m)$ which corresponds to $\zeta = 1$ and $\alpha = \omega$ and implies that in the s-plane all three zeros are superimposed at the point $s = -\alpha_m$. It should be noted that if ξ_3 is disregarded in Eq. (5.5), relationships (ii) and (iii) reduce to Eq. (4.5) which was obtained for the third-order example.

2. Angle of Departure of the Root Locus from the Complex Poles

The use of the definition of the 180° root loci and the geometrical relationships of Fig. 26 leads to the following expression for the angle of departure ϕ:

$$\phi = \theta_1 + \theta_2 + \text{arc tan}\left(\frac{1}{\alpha}\right) - 90 \ . \qquad (5.6)$$

Since θ_1 and θ_2 of Eq. (5.6) are exactly the same angles as θ_1 and θ_2 of Eq. (4.6) in Section IV-B-2, and ξ_1 and ξ_2 are identical in the two cases, it follows that the expression given in Eq. (4.7) is valid here, namely,

$$\theta_1 + \theta_2 = \text{arc tan}\frac{\xi_2}{\xi_1 - 1} \ .$$

By substituting the preceding equation into Eq. (5.6) and replacing α with ξ_3, ϕ is given by

$$\phi = \text{arc tan}\left(\frac{\xi_2}{\xi_1 - 1}\right) + \text{arc tan}\left(\frac{1}{\xi_3}\right) - 90 \ . \qquad (5.7)$$

The equation for the surfaces of constant ϕ in the $\underline{\xi}$ space is

$$\xi_2 = -\left[\frac{1 - \xi_3 \tan(\phi + 90)}{\xi_3 + \tan(\phi + 90)}\right](\xi_1 - 1). \qquad (5.8)$$

Equation (5.8) can be considered as representing a straight line parallel to the ξ_1, ξ_2 plane which intersects the line $\xi_1 = 1$, $\xi_2 = 0$ and has a slope which is dependent upon both ϕ and ξ_3. According to Eq. (5.8), the surface upon which $\phi = 90^\circ$ is expressed by

$$\xi_2 = -\frac{1}{\xi_3}(\xi_1 - 1). \qquad (5.9)$$

It should be noted that when $\xi_3 = 0$ Eq. (5.7) reduces to Eq. (4.7), implying that the intersections of the surfaces of constant ϕ with the ξ_1, ξ_2 plane must be identical with the lines of constant ϕ in Fig. 13b, which were obtained for the third-order example. As $\xi_3 \to \infty$ the lines of constant ϕ parallel to the ξ_1, ξ_2 plane rotate counter-clockwise through 90°.

3. Size of the Dominant Periodic Solution

The results of analog simulations indicate that for the plant under consideration the dominant periodic solution is as shown in Fig. 32, in Appendix A. A condition which the linear switching function must satisfy in order for this periodic solution to exist is derived and the result is expressed in terms of ρ_2 and ρ_4 by Eq. (A6). If ρ_4 is set equal to unity and ρ_2 is expressed in terms of $\underline{\xi}$ by using Eq. (5.4), then Eq. (A6) can be rearranged to yield

$$\xi_2 = -\frac{1}{\xi_3}\left(\xi_1 - \frac{\tan\psi}{\tan\psi - \psi}\right), \quad 0 < \psi \leq \frac{\pi}{2} \tag{5.10}$$

Equation (5.10) states that the size of the dominant periodic solution is constant in the $\underline{\xi}$ space along lines parallel to the ξ_1, ξ_2 plane which intersect the line $\xi_1 = \tan\psi/(\tan\psi - \psi)$, $\xi_2 = 0$ and have a slope of $-1/\xi_3$. When $\psi = \pi/2$ the size of the dominant periodic solution is infinite and Eq. (5.10) reduces to Eq. (5.9), which is the equation for the surface in the $\underline{\xi}$ space along which $\Phi = 90^{\circ}$. For $0 < \psi < \pi/2$, the surfaces of constant periodic solution size are just the surface for $\psi = \pi/2$ shifted out along the ξ_1 axis by the amount $\psi/(\tan\psi - \psi)$.

As in the third-order example, both the nonlinear and the equivalent linear systems give the same boundary in the parameter space for which stability-in-the-large can be inferred.

C. DESIGN OF THE OPTIMAL LINEAR SWITCHING FUNCTION FOR THE FOURTH-ORDER EXAMPLE

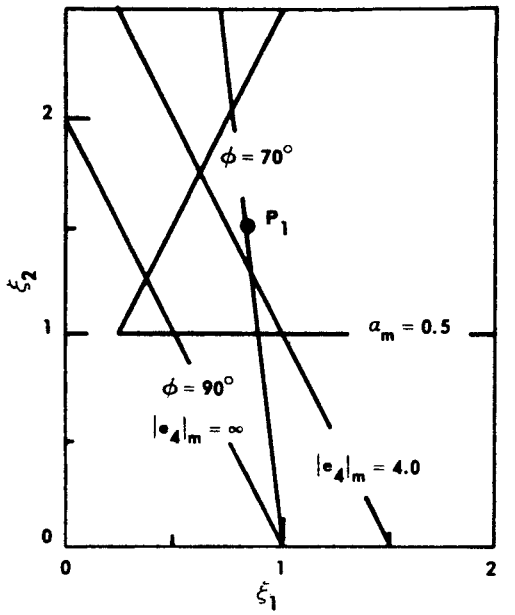
The optimal linear switching function for the plant under consideration and the assumed cost function is found by searching the resultant performance surface to find those values of ρ_1, ρ_2, and ρ_3 which yield the minimum cost. The qualitative design information obtained in the previous section is used to choose several starting points for the random perturbation search process which should lie reasonably close to the optimal parameter values.

The lines of constant value corresponding to $\alpha_m = 0.50$ and 0.707,* $\phi = 70^\circ$ and 90°, and $|e_4|_m = 4.0$ and ∞ are shown in Fig. 27 for $\xi_3 = 1/2$, $1/\sqrt{2}$, and 1. By using this information to apply the three qualitative linear switching-function design criteria, it can be seen that the region of interest in the three-dimensional $\underline{\xi}$ space is most likely within a portion of that shown in Fig. 27. If $\xi_3 < 0.5$ the maximum value of α_m will decrease accordingly, since $\alpha_m \leq \xi_3$, and the small disturbance response of the system in chatter will have a longer time constant than if $\xi_3 \geq 0.5$. Examination of Fig. 27c reveals that when $\xi_3 > 1.0$ stability becomes a problem unless $\alpha_m < 0.5$ because the lines of constant ϕ and constant $|e_4|_m$ rotate counterclockwise as ξ_3 is increased. However, the choice of $\alpha_m < 0.5$ is detrimental to the small-disturbance response, just as when ξ_3 is made small.

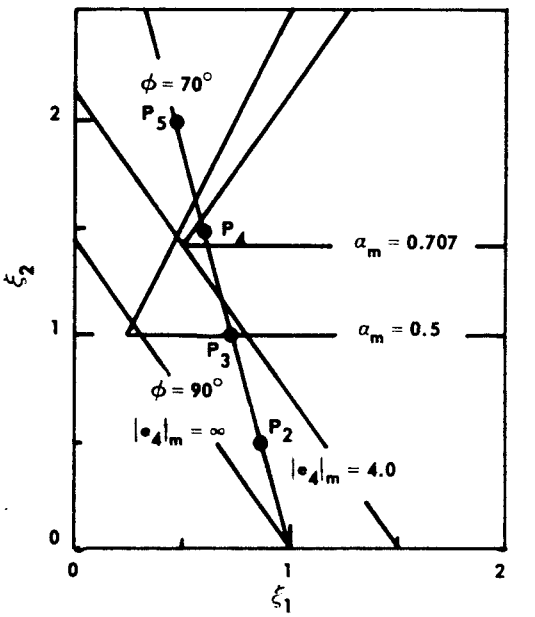
The points denoted P_1, P_3, P_4, and P_6 in Fig. 27 were selected as starting points for the search of the performance surface corresponding to linear switching. All four points satisfy the inequalities $\alpha_m \geq 0.5$ and $\phi \geq 70^\circ$ and $|e_4|_m \simeq 4.0$ in each case. While the choice of these values is somewhat arbitrary, they appear to represent the best compromise between large- and small-disturbance response which can be made with this plant. By the same token, the points denoted P_2 and P_5 in Fig. 27b do not satisfy all three of the above criteria and should, according to the design procedure, yield higher values of cost than the other points selected. In Table 4 the components of $\underline{\xi}$ and $\underline{\rho}$ and the values of J corresponding to these six points are tabulated.

---

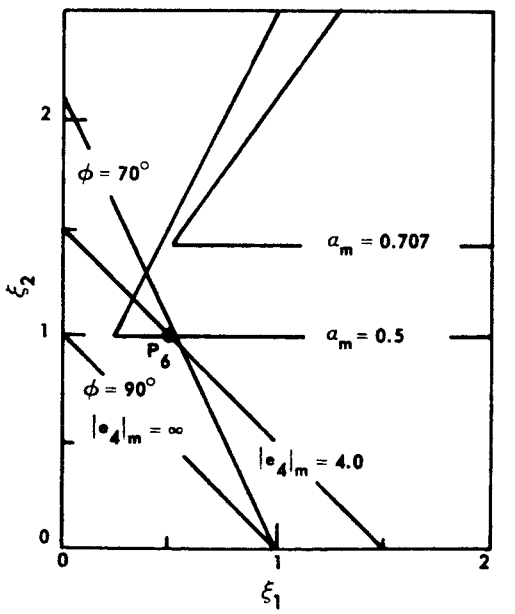
*There is no line for $\alpha_m = 0.707$ in Fig. 27a where $\xi_3 = 0.5$ because Eq. (5.5) requires that $\alpha_m \leq \xi_3$ in order for the plane of constant ξ_3 to intersect the surface of constant minimum real part of the zeros.



a. $\xi_3 = ½$



b. $\xi_3 = 1/\sqrt{2}$



c. $\xi_3 = 1$

FIG. 27. LINES OF CONSTANT LINEAR SWITCHING DESIGN GUIDES IN PLANES OF CONSTANT ξ_3, FOURTH-ORDER EXAMPLE

TABLE 4. VALUES OF $\underline{\xi}$, $\underline{\rho}$ AND J FOR POINTS IN FIG. 27 LABLED P_i, i = 1,...,6

i	ξ_1	ξ_2	ξ_3	ρ_1	ρ_2	ρ_3	ρ_4	J
1	0.82	1.50	0.50	0.41	1.57	2.00	1.00	2.20
2	0.87	0.50	0.707	0.62	1.22	1.21	1.00	2.35
3	0.75	1.00	0.707	0.53	1.46	1.71	1.00	1.45
4	0.63	1.50	0.707	0.45	1.69	2.21	1.00	2.30
5	0.50	2.00	0.707	0.35	1.91	2.71	1.00	3.30
6	0.50	1.00	1.00	0.50	1.50	2.00	1.00	2.20

Examination of Table 4 indicates that, of the four points selected on the basis of the qualitative design procedure (P_1, P_3, P_4, and P_6), all gave finite values for J, i.e., the response was stable for all eight initial conditions used, and one (P_3) yielded a value of cost only 45 percent above the optimal. In addition, of the two points selected as being less satisfactory than these four, one (P_2) yielded a value of J comparable to that obtained with three of them and the other (P_5) yielded a significantly higher cost. Furthermore, several values of $\underline{\xi}$ which were chosen randomly rather than on the basis of the qualitative design criteria either yielded high values of cost (say $J > 3.0$) or resulted in one or more unstable trajectories.

The optimal linear switching function was obtained by starting the three-parameter random search procedure described in Section II-D at those points in the $\underline{\rho}$ space corresponding to points P_3, P_4, and P_6* in the $\underline{\xi}$ space. Of two searches initiated from points P_3 and P_4, one found the absolute minimum and one found a relative minimum. A single search started at point P_6 succeeded in reaching the absolute minimum which was at the point $\hat{\underline{\rho}}^T = (0.72, 1.83, 1.86, 1.00)$** and

---

*Because of the proximity in the $\underline{\rho}$ space of points P_1 and P_6, only P_6 was used as a starting point for the search procedure.

**The value of $\hat{\underline{\xi}}$ is given by $\hat{\underline{\xi}}^T = (1.02, 1.15, 0.71)$, for which $\alpha_m = 0.56$, $\Phi = 54^o$, and $|e_4|_m = 2.7$.

corresponded to a minimum cost of $I = 1.19$. The fact that $\hat{\underline{\rho}}$ was the optimal linear switching function for the specific cost function used was verified by restarting the search procedure at $\hat{\underline{\rho}}$ several times with a relatively large value for the standard deviation of the random parameter perturbations, without improving upon the cost. The step-function responses obtained with the optimal linear switching function are not shown, but they are very similar to those of Fig. 29 which were obtained with PWL switching. Because the corresponding values of cost differ by only 0.07, this similarity is not surprising.

Although the step-function response of the optimal linear switching function is reasonably close to the optimal for $|e^o| \leq 8$, the region of stability is very limited, i.e., $|e^o| > 9$ is unstable. Therefore, in order to increase the region of stability and to possibly lower the minimum cost attainable, the use of PWL switching will be considered.

D. DESIGN OF A PWL SWITCHING FUNCTION FOR THE FOURTH-ORDER EXAMPLE

The value of the four-component vector $\partial\Phi/\partial\underline{\rho}$ in the $\underline{\xi}$ space is used in conjunction with the maximum magnitude attained by the four state-variables e_i during the optimal linear switching responses in order to select the switching function component(s) to be made PWL. The maximum magnitudes attained by e_1, e_2, e_3, and e_4 were observed to be 8.0, 3.0, 1.5, and 2.4, respectively. These values are used in guiding the selection of the breakpoint(s). The vector $\partial\Phi/\partial\underline{\rho}$ is evaluated by the straightforward application of Eq. (3.12) to the expression for $\Phi(\underline{\xi})$ given by Eq. (5.7). After a considerable amount of algebra it is found that

$$\frac{\partial\Phi}{\partial\rho_1} = -\left\{\frac{(\xi_1 - 1) + \xi_2\xi_3}{(1 + \xi_3^2)\left[(\xi_1 - 1)^2 + \xi_2^2\right]}\right\}, \qquad (5.11a)$$

$$\frac{\partial\Phi}{\partial\rho_2} = \left\{\frac{\xi_3(\xi_1 - 1) - \xi_2}{(1 + \xi_3^2)\left[(\xi_1 - 1)^2 + \xi_2^2\right]}\right\}, \qquad (5.11b)$$

$$\frac{\partial \phi}{\partial \rho_3} = - \frac{\partial \phi}{\partial \rho_1} \quad , \tag{5.11c}$$

and

$$\frac{\partial \phi}{\partial \rho_4} = - \frac{\partial \phi}{\partial \rho_2} \quad . \tag{5.11d}$$

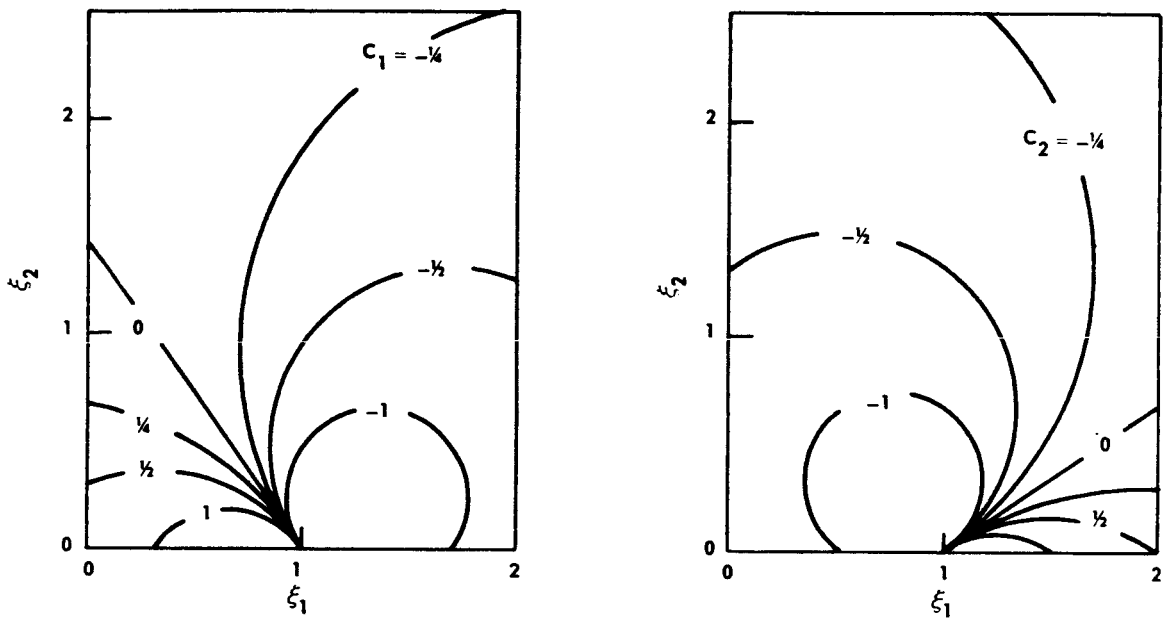
The intersections of the surfaces of constant $\partial\phi/\partial\underline{\rho}$ with the plane $\xi_3 = 1/\sqrt{2}$ are shown in Fig. 28.

Since e_1 varies over a substantially wider range than do the other components of $\underline{e}$, it is natural to consider making σ_1 PWL. If a single breakpoint is used initially, the parameter array $\{\rho\}$ has five undetermined components and appears as

$$\{\rho\} = \begin{Bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & x & x \\ \rho_{31} & x & x \\ 1 & x & x \end{Bmatrix} \tag{5.12}$$

where the x's denote elements which are undefined for this particular choice of switching function. Examination of Eq. (5.11a) shows that $\partial\phi/\partial\rho_1 = 0$ along the surface defined by $\xi_2 = - \frac{1}{\xi_3} (\xi_1 - 1)$ which, according to Eq. (5.9), is the same surface along which $\phi = 90^o$. Therefore, if any improvement in the large disturbance response is to be obtained by making σ_1 be PWL, it is necessary that the value of ϕ corresponding to the linear portion of the switching function be somewhat less than 90^o. Furthermore, since $\partial\phi/\partial\rho_1 < 0$ when $\phi < 90^o$, the equivalent slope of σ_1 must be decreased beyond the breakpoint in order to raise the effective value of ϕ, i.e., $\rho_{13} < \rho_{11}$.

The search of the performance surface corresponding to the array (5.12) was carried out on the hybrid computer by fixing the breakpoint ρ_{12} and the corresponding change of slope $(\rho_{13} - \rho_{11})$ and optimizing the three linear switching parameters ρ_{11}, ρ_{21}, and ρ_{31} with the



a. $\partial\phi/\partial\rho_1 = C_1 = -\partial\phi/\partial\rho_3$

b. $\partial\phi/\partial\rho_2 = C_2 = -\partial\phi/\partial\rho_4$

FIG. 28. LINES OF CONSTANT $\partial\phi/\partial\underline{\rho}$ IN THE PLANE $\xi_3 = 1/\sqrt{2}$, FOURTH-ORDER EXAMPLE

random search procedure. Initially ρ_{11}, ρ_{21}, and ρ_{31} were set at the values yielding the optimal linear switching function and the PWL parameters were chosen to be $\rho_{12} = 6.0$ and $(\rho_{13} - \rho_{11}) = -0.25$ in order to make the departure from the optimal linear switching function small. After the three linear switching parameters were optimized for several values of $(\rho_{13} - \rho_{11})$ the breakpoint was lowered and the process repeated, but with the linear switching parameters starting from their current optimum values using the PWL switching function. In this manner, minimum costs of $J = 1.17$, 1.12, 1.13, and 1.18 were obtained with $\rho_{12} = 6.0$, 5.0, 4.0, and 3.0, respectively.* The optimal parameter values yielding $J = 1.12$ were given by the array

$$\{\rho\} = \begin{Bmatrix} 0.98 & 5.00 & 0.48 \\ 2.18 & x & x \\ 2.00 & x & x \\ 1.00 & x & x \end{Bmatrix} \qquad (5.13)$$

where the x's denote undefined elements, and the corresponding step-function responses are shown in Fig. 29.

An examination of trajectories with non-zero values for e_2^o, e_3^o, and e_4^o reveals that, as with the optimal linear switching function, the region of stability in the state-space is severely limited. For example, when $e_2^o = e_4^o = 0$, the maximum value of e_3^o for which the response is stable diminishes from 1.60 when $e_1^o = 0$ to only 0.45 when $e_1^o = 10$. In Appendix A it is shown that the form of the equation for the locus of possible switching points for the dominant periodic solution is identical to that for the $1/s(s^2 + 1)$ plant, but with the e_2 and e_4 coordinates substituted for the e_1 and e_3 coordinates, respectively, and with $e_1 = e_3 = 0$. The switching surface corresponding

---

*The minimum value of J attainable with linear switching was $I = 1.19$.

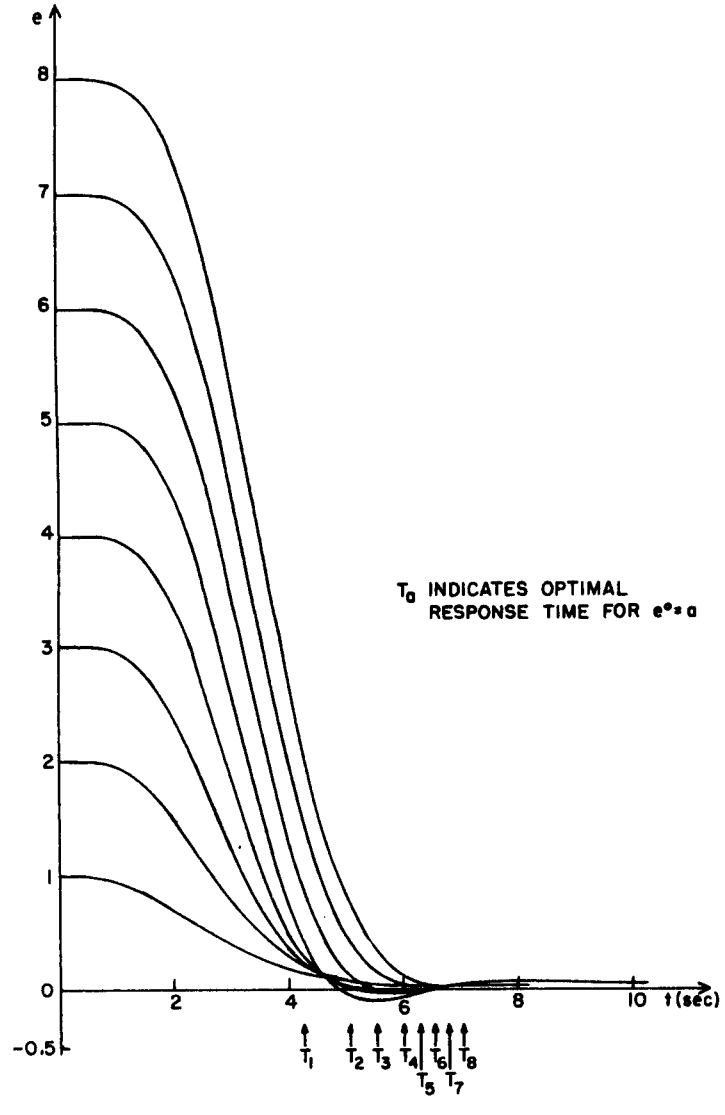


FIG. 29. STEP-FUNCTION RESPONSES WITH THE PIECE-WISE-LINEAR SWITCHING FUNCTION CORRESPONDING TO THE ARRAY (5.13), FOURTH-ORDER EXAMPLE

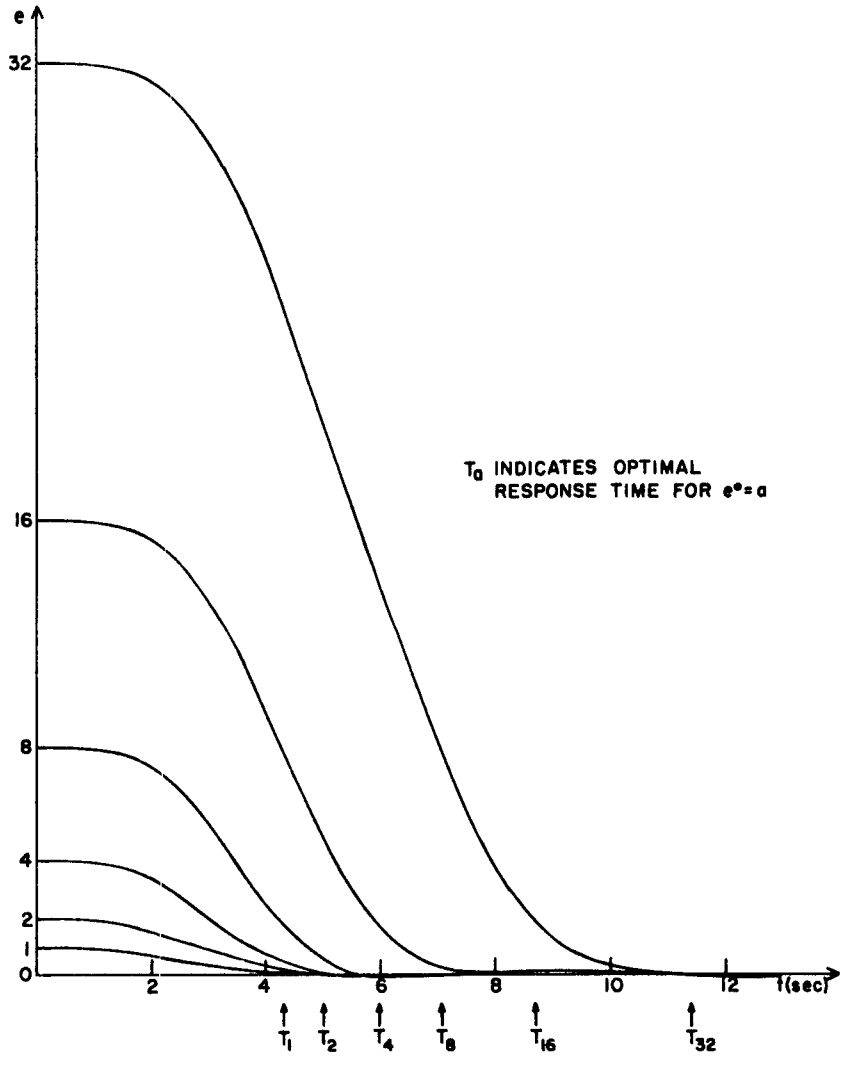


FIG. 30. STEP-FUNCTION RESPONSES WITH THE PIECE-WISE-LINEAR SWITCHING FUNCTION CORRESPONDING TO THE ARRAY (5.15), FOURTH-ORDER EXAMPLE

to the array (5.13) intersects the e_2, e_4 plane along the line $e_4 = -2.18e_2$ and the necessary condition for the existence of the dominant periodic solution is satisfied. To make it impossible for this periodic solution to exist σ_4 was made PWL with a breakpoint at $\rho_{42} = 1.7$ and with $\rho_{43} = 2.40$, yielding an intersection of the switching surface and the e_2, e_4 plane which no longer intersects the locus of possible periodic solution switching points, thereby making it impossible for the dominant periodic solution to exist. Making σ_4 PWL affected several of the trajectories used to measure the cost corresponding to the array (5.13). Therefore, the three linear switching parameters were adjusted slightly, using the random search procedure, to yield the array

$$\{\rho\} = \begin{Bmatrix} 0.96 & 5.00 & 0.46 \\ 2.26 & x & x \\ 2.00 & x & x \\ 1.00 & 1.70 & 2.40 \end{Bmatrix} \qquad (5.14)$$

for which $J = 1.15$. It was verified on the analog computer that the origin was asymptotically stable for all initial conditions which could be simulated, namely $\|\underline{e}\| \leq 100$.

One of the primary reasons for using a PWL switching function as opposed to a linear switching function is to extend the range of initial conditions for which satisfactory performance can be achieved. To illustrate the usefulness of a PWL switching function when the step-function magnitude varies over the range $1 \leq |e^o| \leq 32$, the switching function parameters

$$\{\rho\} = \begin{Bmatrix} 0.70 & 8.00 & 0.50 & 16.00 & 0.05 \\ 1.73 & x & x & x & x \\ 1.65 & x & x & x & x \\ 1.00 & x & x & x & x \end{Bmatrix} \qquad (5.15)$$

were obtained by observing the trajectories with the analog computer in the repetitive mode and adjusting the parameters manually. The corresponding transient responses are shown in Fig. 30 for several representative values of e^{o}. As such, this performance represents an initial parameter set from which an optimization could be started if the performance is not adequate as given.

It has been possible to obtain an optimal PWL switching function for a dominant fourth-order plant. However, it should be emphasized that the task of searching a high-dimensionality surface, while presenting no formal difficulty, may represent a substantial challenge to the designer. Therefore, when the plant is truly of fifth- or higher-order, i.e., is not adequately represented by a plant of lower-order, it is felt that careful consideration must be given to the choice of the surface searching procedure to be employed and, as previously mentioned, the methods described by Kushner [Ref. 12] and Brown [Ref. 13] merit consideration, provided that sufficient digital computer capacity is available.

VI. CONCLUSION

A. SUMMARY

The problem considered in this dissertation is the possible use of PWL switching functions for satisfying the minimum-time criterion and for that class of control systems which can be represented by a linear, constant parameter plant whose transfer function has only poles, a single controller which is an ideal contactor, and an input function which is equivalent to initial conditions on the state-variables. Furthermore, it has been assumed that all state-variables are observable and that no random effects exist.

In Chapter II a performance criterion is defined in terms of the weighted response times for an unspecified number of initial conditions. The effect of the number of initial conditions upon the resulting performance surface is investigated in detail for the $1/s^2$ plant. It is shown qualitatively that as more initial conditions are used in the definition of the cost the performance surface becomes more amenable to a surface searching process in the sense that the number and size of relative minima are reduced, fluctuations in the surface gradient are lessened, and the region near the absolute minimum becomes more nearly parabolic. Also, it is shown, again for the second-order example, that the minimum cost obtainable is relatively insensitive to the locations of the breakpoints of the PWL function provided that the corresponding slopes are adjusted to their optimum values. Two surface searching methods used are presented and the results obtained with the $1/s^2$ plant are described.

In Chapter III a qualitative method based upon root-locus techniques and the existence of certain periodic solutions is given for designing sub-optimal linear switching functions. It is shown how the properties of the root loci and the periodic solution can be used to provide qualitative information as to which of the linear switching function components should be made PWL functions and in what directions the slopes should be changed. A second-order plant with unstable oscillatory poles is used to illustrate the qualitative design procedure for both linear and PWL switching functions.

The design of an optimal PWL switching function for the third-order plant $1/s(s^2 + 1)$ subject to step-function inputs is carried out in detail in Chapter IV. It is shown that the results of the qualitative design procedure are in good agreement with the measured linear switching performance surface corresponding to the assumed cost function. By using the root-locus information to select a linear switching-function component to be made PWL and then using the random search procedure to optimize the parameter values, a PWL switching function having a single breakpoint is found which yields close-to-optimal cost. By considering the existence of the dominant periodic solution, it is shown how the PWL switching function designed on the basis of cost alone can be modified so that the effects of parameter variations upon the system stability are substantially reduced and the region of stability in the state-space is increased by a factor of at least one hundred. The third-order plant $1/(s + 0.5)(s^2 + 0.8s + 1)$ is considered briefly to show that there is no difficulty in applying the design methods to plants with well-damped poles. It is found that, while linear switching gives close-to-optimal responses for large initial conditions for such well-damped plants, PWL switching gives close-to-optimal responses for both large and small initial conditions. It is indicated in Chapter V that there are no theoretical limits to the order of the plant which can be treated by the proposed methods, and this is illustrated by designing a PWL switching function for the plant $1/s^2(s^2 + 1)$. However, it is evident that the practical aspects of carrying out the performance surface search do become more complex as the plant order increases.

This investigation has provided three contributions to the field of contactor control system technology. First, it has been demonstrated quantitatively that relatively simple PWL switching functions which, by their very nature, are easily implemented can give very close-to-minimum-time responses for a variety of plants and initial conditions. In addition, PWL switching functions can be used to enlarge greatly the region of stability in the state-space for those systems with linear or PWL switching having finite stable regions.

Second, a procedure has been given for qualitatively designing linear switching functions for plants of arbitrary order and for determining which switching function components should be made PWL and how the slopes should be changed in order to improve the performance over that attainable with linear switching.

Finally, a procedure has been presented for the synthesis of PWL switching functions by formulating the problem as the search of a performance surface and by providing methods for defining a performance surface which will be amenable to standard searching techniques, for initiating the search, and for reducing the dimensionality of the surface.

The principal advantages of PWL switching functions and of the proposed methods are:

1. PWL switching functions of the type considered here are extremely simple to implement and are suitable for use in adaptive systems.

2. The design procedures are not limited theoretically by the order of the plant.

3. No a priori assumptions are made regarding the number of switching points or the shape of the optimal switching surface, so presumably PWL switching and the design methods are applicable to a wide variety of plants and initial conditions.

4. A quantitative performance criterion is used for the design of optimal linear and PWL switching functions.

Likewise, the following practical considerations should be pointed out:

1. If the optimization procedure is performed entirely on a digital computer, as opposed to using a hybrid system, a considerable amount of computation time is required, primarily for the integration of differential equations.

2. As the order of the plant and, in turn, the dimensionality of the performance surface increase, the task of finding the optimal parameter values presents an increasingly complex surface searching problem.

Although the usefulness of PWL switching and of the design procedures presented must be demonstrated with specific plants, initial conditions, and cost functions, it is possible to draw several general conclusions regarding their applicability on the basis of the results presented here. The design procedures developed here appear to be applicable to plants having their poles anywhere in the s-plane except possibly on or close to the positive real axis. Higdon [Ref. 21] has investigated plants with unstable real roots from a different point of view and has shown that PWL switching surfaces are of considerable utility in that case. In general, PWL switching provides improved performance over that attainable with linear switching when the plant poles are not well-damped and when the initial conditions being considered do not have approximately co-planar optimal switching points. When the plant poles are well-damped the cost-free region must be relatively small in order to warrant PWL switching.

It is felt that PWL switching is useful in an engineering sense for a larger variety of plants and initial conditions than any single one of the quasi-optimal design methods discussed in Section I-B. Furthermore, it is apparent that the simplicity of the resulting system makes PWL switching very attractive for practical applications.

B. SUGGESTIONS FOR FURTHER INVESTIGATION

Since the range of applicability of the class of PWL switching functions considered here can be determined only by direct verification, it would be useful to consider examples for a wider variety of plants and initial conditions than was possible to evaluate here, and, in particular, for unstable plants. Also, methods to reduce the complexity of the surface searching problem when the order of the plant is high are worthy of consideration.

It seems likely that a systematic method of using some of the initial conditions to determine the linear portion of the switching surface and then using the remainder to determine the PWL parameters could be devised. The problems of selecting which initial conditions are to be used in determining the respective switching function parameters and of insuring

that the trajectories for large disturbances can be made stable once the linear portion of the switching function is fixed do not appear to be trivial.

The questions raised by the observed influence of the so-called dominant periodic solution upon the region of stability in a three- or higher-dimensional state-space should be answered. In particular, the relationship between the observed dominant periodic solution and the stability boundary in the n-dimensional state-space, and the existence and meaning of other periodic solutions are topics whose understanding may very well lead to useful results in an area which has not received a great deal of attention to date.

Finally, the ever-present question of equivalent-linearization is in need of rigorous resolution. Although Aizerman's conjecture is invalid in a number of cases, it is likely that useful results could be obtained by restricting the nonlinearity to being an ideal contactor rather than allowing it to have the very general nature for which the conjecture was made.

APPENDIX A. A CONDITION NECESSARY FOR THE EXISTENCE OF THE DOMINANT PERIODIC SOLUTION FOR THE $1/s(s^2+1)$ AND $1/s^2(s^2+1)$ PLANTS

In this appendix a set of conditions necessary for the existence of the observed dominant periodic solution will be given for the plants considered in Chapters IV and V. No attempt will be made to consider sufficient conditions or other types of periodic solutions.

1. $1/s(s^2+1)$ PLANT

A typical dominant periodic solution which has been observed for this plant is as shown by Fig. 31. Examination of the projections of the trajectory indicate that the coordinates of the two switching points N_1 and N_2 in the canonical space are given by

$$\underline{x}^T = \pm (\psi, 0, \tan \psi) \qquad 0 < \psi < \pi/2.$$

By using the transformation of Eq. (4.3b), the coordinates of the switching points in the error space, expressed in terms of the quarter-period ψ, are found to be

$$\underline{e}^T = \pm (\psi - \tan \psi, 0, \tan \psi) \; . \tag{A1}$$

Since the two points N_1 and N_2 are to be switching points it is necessary that the switching function be zero at these points if the dominant periodic solution is to exist. If a linear switching function is used and ρ_3, the coefficient of e_3, is not set equal to unity for the moment, it is seen that at the points N_1 and N_2

$$\rho_1 e_1 + \rho_2 e_2 + \rho_3 e_3 = 0. \tag{A2}$$

Using Eq. (A1) to obtain the coordinates of N_1 and N_2, it follows that the condition necessary for the existence of the dominant periodic solution is that a solution exist to the equation

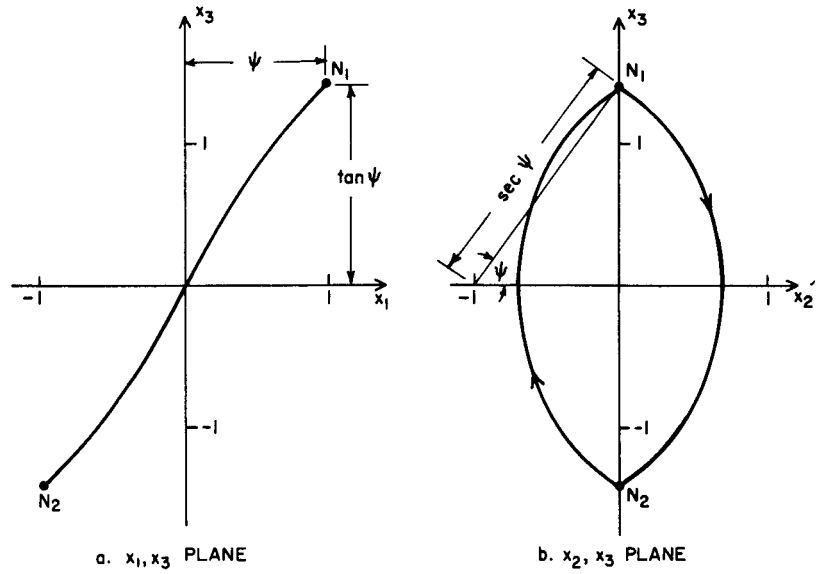


FIG. 31. TYPICAL DOMINANT PERIODIC SOLUTION FOR THE $1/s(s^2+1)$ PLANT

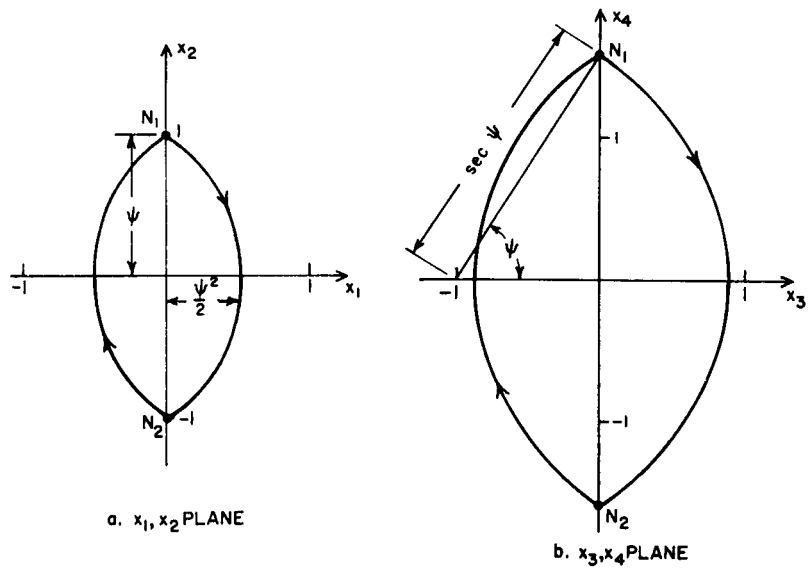


FIG. 32. TYPICAL DOMINANT PERIODIC SOLUTION FOR THE $1/s^2(s^2+1)$ PLANT

$$\frac{\tan\psi}{\psi} = \frac{\rho_1}{\rho_1 - \rho_3} \qquad 0 < \psi < \pi/2. \tag{A3}$$

Since both ρ_1 and ρ_3 must be positive to insure stable chatter, Eq. (A3) can have a solution if and only if

$$\rho_1 > \rho_3 .$$

It is also of interest to consider a graphical derivation of the preceding inequality. If the parameter ψ is eliminated from Eq. (A1), a locus of points lying in the e_1, e_3 plane is obtained, obeying the equation

$$e_1 = - e_3 + \arctan e_3 , \qquad e_2 = 0. \tag{A4}$$

This line in the e_1, e_3 plane, which is plotted in Fig. 33, represents all points in the state-space which can be switching points for the particular periodic solution depicted in Fig. 31. For large values of $|e_3|$ the locus of possible switching points is asymptotic to the lines $e_1 = -e_3 + \pi/2 \operatorname{sgn} e_3$ which have a slope of -1 in the e_1, e_3 plane. The intersection of the linear switching plane described by Eq. (A2) with the e_1, e_3 plane is the line

$$e_3 = - \frac{\rho_1}{\rho_3} e_3 . \tag{A5}$$

If the slope of this line is greater than -1, i.e., $\rho_1 < \rho_3$, there can be no non-trivial intersection between the line described by Eq. (A5) and the locus of possible switching points described by Eq. (A4). Two typical switching surface intersections are shown in Fig. 33. Line A is the switching plane intersection corresponding to $\rho_1/\rho_3 = 3.2$ and, as verified by analog simulation, produces the dominant periodic solution shown in Fig. 31 when ρ_2 is chosen in the range $0.5 < \rho_2 < 3$. Line B corresponds to $\rho_1/\rho_3 = 0.5$ and indicates that the dominant periodic solution cannot exist, as verified by analog simulation.

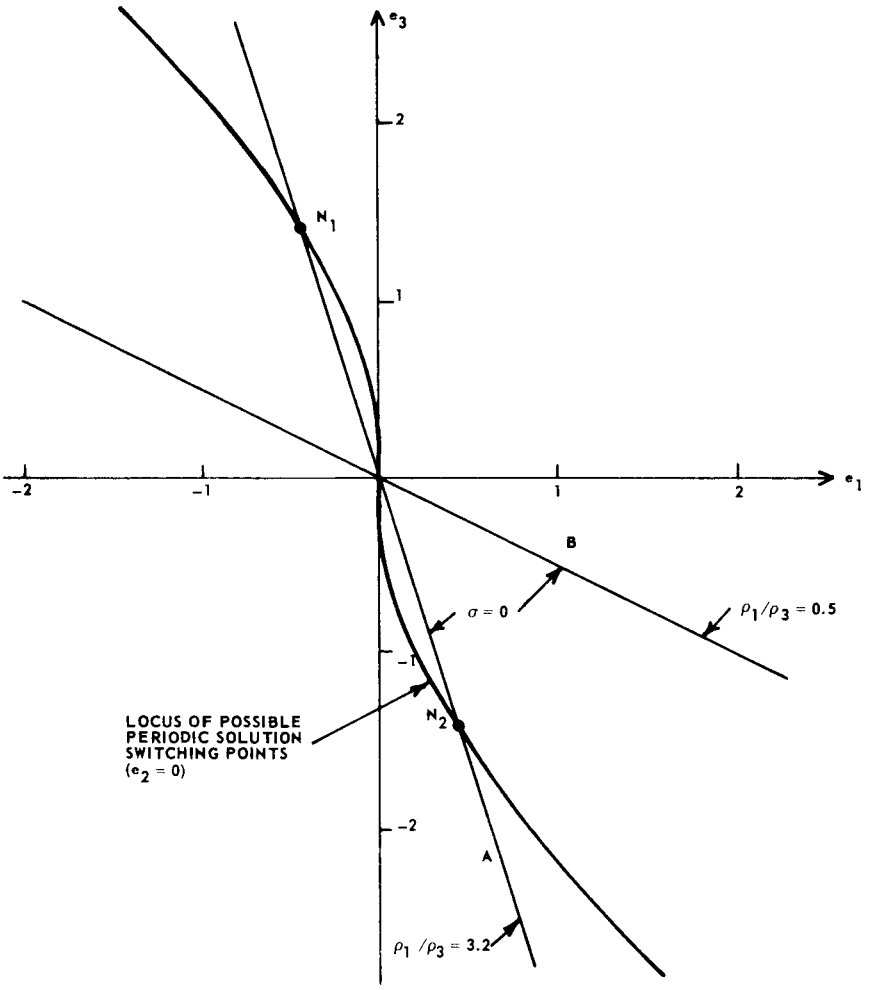


FIG. 33. LOCUS OF POSSIBLE PERIODIC SOLUTION SWITCHING POINTS AND INTERSECTIONS OF TWO LINEAR SWITCHING SURFACES WITH THE e_1, e_3 PLANE, $1/s(s^2+1)$ PLANT

2. $1/s^2(s^2+1)$ PLANT

The observed dominant periodic solution for this plant is of the form shown by Fig. 32. By inspection it can be seen that the coordinates of the two switching points N_1 and N_2 in the canonical space are given by

$$\underline{x}^T = \pm\ (0,\ \psi,\ 0,\ \tan\psi) \qquad 0 < \psi < \pi/2.$$

By application of the transformation of Eq. (5.2b) the coordinates of N_1 and N_2 in the error space are found to be

$$\underline{e}^T = \pm\ (0,\ \psi - \tan\psi,\ 0,\ \tan\psi)\ .$$

If the two points given by this equation are to lie on the linear switching surface, it is necessary that

$$\frac{\tan\psi}{\psi} = \frac{\rho_2}{\rho_2 - \rho_4} \qquad 0 < \psi < \pi/2. \tag{A6}$$

Therefore, the condition necessary for the existence of the dominant periodic solution is

$$\rho_2 > \rho_4\ .$$

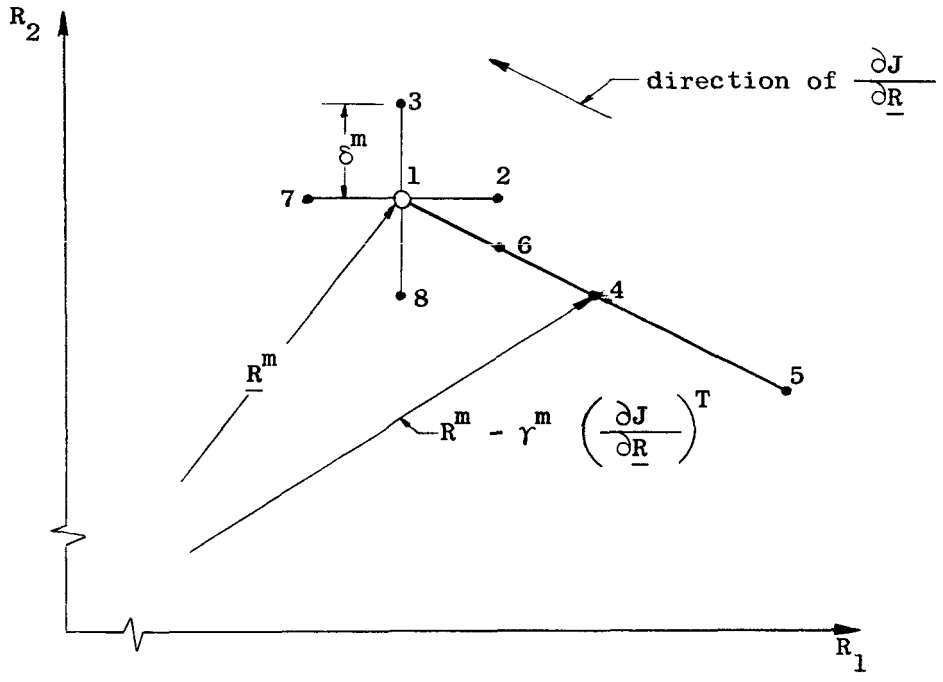
As in the third-order example of the previous section, this inequality can be shown graphically by eliminating ψ from Eq. (A6) to obtain the equation for the locus of possible periodic solution switching points:

$$e_2 = -e_4 + \arctan e_4\ , \qquad e_1 = e_3 = 0.$$

Because the form of the preceding equation is identical to that of Eq. (A4), it follows that the locus of possible periodic solution switching points is exactly as shown in Fig. 33, except with e_2 and e_4 substituted for e_1 and e_3, respectively, and with $e_1 = e_3 = 0$.

APPENDIX B. A TWO-PARAMETER GRADIENT SEARCH PROCESS

To illustrate the manner in which the gradient surface searching procedure mentioned in Section II-E is accomplished, consider the following sketch where the points corresponding to possible measurements of the cost J in the two-dimensional R_1, R_2 parameter space during the m^{th} iteration cycle are shown.



Point 1 denotes the value of $\underline{R}$ at the start of the m^{th} iteration and the corresponding value of the cost function is J_1. First, the two components of the approximate surface gradient are measured according to the relationships

$$\frac{\partial J}{\partial R_1} \doteq \frac{J_2 - J_1}{\delta^m} \quad \text{and} \quad \frac{\partial J}{\partial R_2} = \frac{J_3 - J_1}{\delta^m} ,$$

where J_2 and J_3 denote the values of J measured at points 2 and 3 respectively; the parameter δ^m is the magnitude of the perturbation made in the components of $\underline{R}^m$ in order to measure the approximate gradient at point 1. The value of $\underline{R}$ corresponding to point 4 is then computed using Eq. (2.6) and the cost J_4 can then be measured.

If $J_4 \leq J_1$, the change in $\underline{R}$ previously computed by using Eq. (2.6) is then doubled, yielding point 5, and J_5 is measured. If $J_5 \leq J_4$, point 5 is taken as $\underline{R}^{m+1}$ and the process is repeated to find $\underline{R}^{m+2}$. If $J_4 > J_1$, the computed change in $\underline{R}$ is halved, yielding point 6 and the corresponding J_6. If $J_6 \leq J_1$, point 6 is taken as $\underline{R}^{m+1}$ and the process is repeated. By using this procedure, the effective value of the feedback gain γ can vary from one-half to twice its nominal value, with only a 25 percent increase in the number of points at which J must be measured.

If both J_4 and J_6 are greater than J_1, either the measurement of the approximate gradient has furnished misleading information or the value of γ^m was too high, provided, of course, that noise has not been a significant factor. Therefore, it would seem reasonable to reduce γ and δ and repeat the process. However, before doing this the values of cost used in measuring the gradient, namely J_2 and J_3 are compared with J_1 and, if either one is less than or equal to J_1, the corresponding value of $\underline{R}$ is taken as $\underline{R}^{m+1}$. If neither of these tests provides a lower value of cost, then J_7 and J_8 are measured and compared to J_1. Finally, if both J_7 and J_8 exceed J_1, the values of γ and δ are halved and the process repeated, again starting with the measurement of the approximate gradient at point 1. The search process is stopped when δ has reached a certain arbitrary lower bound.

If the trajectory corresponding to any one of the K initial conditions used in measuring J is unstable or has a settling time which exceeds the limits for which the analog computer is scaled, the value of J at that particular $\underline{R}$ cannot be determined, i.e., the performance surface does not exist there. The procedure followed in this event depends upon the particular point in the search process at which the instability occurs. If J_1 does not exist for the first iteration,

the search process cannot be initiated and a new choice of $\underline{R}$ must be made. Methods for initiating the search process at a point in the parameter space where stability can be guaranteed are discussed in Chapter III. If J_1 exists but either J_2 or J_3 does not, the value of δ must be reduced until both J_2 and J_3 exist before the search process can continue. Since, in this situation, the value of $\underline{R}$ lies near a region where $J = \infty$, the value of $\partial J/\partial \underline{R}$ is likely to be quite large, so the value of γ is also halved. If J_4 does not exist, then J_6 is measured. If J_6 does not exist either, both γ and δ are halved. Should J_5 be infinite, $\underline{R}^{m+1}$ is taken as point 4 because J_5 may be measured only if $J_4 \leq J_1$. Finally, if either J_7 or J_8 are measured and found to be infinite, both γ and δ are halved and the approximate gradient is remeasured.

The choice of the initial values for γ and δ was made by trial and error, but, because of their adaptive nature, the choice was not critical in the examples worked, provided that they were not chosen too low.

REFERENCES

1. L. S. Pontryagin et al, The Mathematical Theory of Optimal Processes, John Wiley and Sons, Inc., New York, 1962.

2. R. A. Hubbs, "Adaptive Quasi-Optimum Switching Surfaces for Contactor Control Systems," TR No. 2106-4, Stanford Electronics Laboratories, Stanford, Calif., April 1962.

3. S. F. Schmidt, "The Analysis and Design of Continuous and Sampled-Data Feedback Control Systems with a Saturation Type Nonlinearity," NASA Tech. Note D-20, Ames Research Center, Moffett Field, Calif., Aug. 1959.

4. S. F. Schmidt and E. V. Harper, "The Design of Feedback Control Systems Containing a Saturation Type Nonlinearity," NASA Tech. Note D-324, Ames Research Center, Moffett Field, Calif., Sept. 1960.

5. A. A. Feld'baum, Rechengeräte in Automatischen Systemen, R. Oldenbourg Verlag, Munich, 1962.

6. I. Flügge-Lotz and H. A. Titus, Jr., "Optimum and Quasi-Optimum Control of Third and Fourth-Order Systems," Tech. Report No. 134, Division of Engineering Mechanics, Stanford University, Stanford, Calif., Oct. 1962, (an abbreviated version is to be published in Proc. Second International Congress of Automatic Control, Butterworth's, 1963).

7. I. Flügge-Lotz and H. E. Lindberg, "Studies of Second and Third-Order Contactor Control Systems," NASA Tech. Note D-107, Stanford University, Stanford, Calif., 1959.

8. I. Flügge-Lotz, Discontinuous Automatic Control, Princeton University Press, Princeton, N. J., 1953.

9. I. Flügge-Lotz and T. Ishikawa, "Investigation of Third-Order Contactor Control Systems with Two Complex Poles Without Zeros," NASA Tech. Note D-428, Stanford University, Stanford, Calif., June 1960.

10. I. Flügge-Lotz, "Synthesis of Third-Order Contactor Control Systems," Proc. First International Congress on Automatic Control (Moscow 1960), Butterworth's, 1960.

11. C. E. Hutchinson, "Minimum-time Control of a Linear Combination of State Variables," TR No. 6311-1, Stanford Electronics Laboratories, Stanford, Calif., Aug. 1963.

12. H. J. Kushner, "Hill Climbing Methods for the Optimization of Multi-parameter Noise Disturbed Systems," Trans. ASME, Journal of Basic Engineering, June 1963, pp. 157-164.

13. R. R. Brown, "Gradient Methods for the Computer Solution of System Optimization Problems," MIT WADC Tech. Note 57-159, Sept. 1957.

14. R. W. Hamming, Numerical Methods for Scientists and Engineers, McGraw-Hill Book Co., Inc., New York, 1962, p. 389.

15. R. E. Kalman, "Physical and Mathematical Mechanisms of Instability in Nonlinear Automatic Control Systems," Trans. ASME, 79, 1957, pp. 553-566.

16. W. Hahn, Theory and Application of Liapunov's Direct Method, Prentice-Hall, Inc., Englewood Cliffs, N. J., 1963.

17. J. Gille, M. Pelegrin, and P. Decaulne, Feedback Control Systems, McGraw-Hill Book Co., Inc., New York, 1959.

18. I. Flügge-Lotz and M. Yin, "On the Optimum Response of Third-Order Contactor Control Systems," Tech. Report No. 125, Division of Engineering Mechanics, Stanford University, Stanford, Calif., April 1960, (an abbreviated version appears in Trans. ASME, Journal of Basic Engineering, 1961, pp. 59-64).

19. D. Bushaw, "Optimal Discontinuous Forcing Terms," Contributions to the Theory on Nonlinear Oscillations, IV, Ed. by S. Lefschetz, Princeton University Press, Princeton, N. J., 1958.

20. Y. Kashiwagi, Division of Engineering Mechanics, Stanford University, Stanford Calif., unpublished results, 1963, (see discussion to [Ref. 6] in Proc. Second International Congress on Automatic Control).

21. D. Higdon, "Automatic Control of Inherently Unstable Dynamic Systems with Bounded Control Inputs," Tech. Report No. 176, Division of Aeronautics and Astronautics, Stanford University, Stanford, Calif., 1963.